ALGEBRA

Michael Artin
Massachusetts Institute of Technology









PRENTICE HALL
Upper Saddle River, New Jersey 07458

Library of Congress Cataloging-in-Publication Data

Artin, Michael.
Algebra / Michael Artin.
p. cm.
Includes bibliographical references and index.
ISBN 0-13-004763-5
1. Algebra. I. Title.
QA154.2.A77 1991 91-2107
512.9—dc20 CIP

Figure 4.16 from *Zeitschrift für Kristallographie*

Editorial/Production Supervision and
Interior Design: Ruth Cottrell
Prepress Buyer: Paula Massenaro
Manufacturing Buyer: Lori Bulwin



A Simon & Schuster Company
Upper Saddle River, New Jersey 07458



Printed in the United States of America
10 9 8

ISBN 0-13-004763-5

Prentice-Hall International (UK) Limited, *London*
Prentice-Hall of Australia Pty. Limited, *Sydney*
Prentice-Hall Canada Inc., *Toronto*
Prentice-Hall Hispanoamericana, S.A., *Mexico*
Prentice-Hall of India Private Limited, *New Delhi*
Prentice-Hall of Japan, Inc., *Tokyo*
Simon & Schuster Asia Pte. Ltd., *Singapore*
Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*

To my wife Jean

Contents

Preface

Important though the general concepts and propositions may be with which the modern and industrious passion for axiomatizing and generalizing has presented us, in algebra perhaps more than anywhere else, nevertheless I am convinced that the special problems in all their complexity constitute the stock and core of mathematics, and that to master their difficulties requires on the whole the harder labor.

Herman Weyl

This book began about 20 years ago in the form of supplementary notes for my algebra classes. I wanted to discuss some concrete topics such as symmetry, linear groups, and quadratic number fields in more detail than the text provided, and to shift the emphasis in group theory from permutation groups to matrix groups. Lattices, another recurring theme, appeared spontaneously. My hope was that the concrete material would interest the students and that it would make the abstractions more understandable, in short, that they could get farther by learning both at the same time. This worked pretty well. It took me quite a while to decide what I wanted to put in, but I gradually handed out more notes and eventually began teaching from them without another text. This method produced a book which is, I think, somewhat different from existing ones. However, the problems I encountered while fitting the parts together caused me many headaches, so I can't recommend starting this way.

The main novel feature of the book is its increased emphasis on special topics. They tended to expand each time the sections were rewritten, because I noticed over the years that, with concrete mathematics in contrast to abstract concepts, students often prefer more to less. As a result, the ones mentioned above have become major parts of the book. There are also several unusual short subjects, such as the Todd–Coxeter algorithm and the simplicity of PSL_2.

In writing the book, I tried to follow these principles:

1. The main examples should precede the abstract definitions.
2. The book is not intended for a "service course," so technical points should be presented only if they are needed in the book.
3. All topics discussed should be important for the average mathematician.

Though these principles may sound like motherhood and the flag, I found it useful to have them enunciated, and to keep in mind that "Do it the way you were taught" isn't one of them. They are, of course, violated here and there.

The table of contents gives a good idea of the subject matter, except that a first glance may lead you to believe that the book contains all of the standard material in a beginning algebra course, and more. Looking more closely, you will find that things have been pared down here and there to make space for the special topics. I used the above principles as a guide. Thus having the main examples in hand before proceeding to the abstract material allowed some abstractions to be treated more concisely. I was also able to shorten a few discussions by deferring them until the students have already overcome their inherent conceptual difficulties. The discussion of Peano's axioms in Chapter 10, for example, has been cut to two pages. Though the treatment given there is very incomplete, my experience is that it suffices to give the students the flavor of the axiomatic development of integer arithmetic. A more extensive discussion would be required if it were placed earlier in the book, and the time required for this wouldn't be well spent. Sometimes the exercise of deferring material showed that it could be deferred forever—that it was not essential. This happened with dual spaces and multilinear algebra, for example, which wound up on the floor as a consequence of the second principle. With a few concepts, such as the minimal polynomial, I ended up believing that their main purpose in introductory algebra books has been to provide a convenient source of exercises.

The chapters are organized following the order in which I usually teach a course, with linear algebra, group theory, and geometry making up the first semester. Rings are first introduced in Chapter 10, though that chapter is logically independent of many earlier ones. I use this unusual arrangement because I want to emphasize the connections of algebra with geometry at the start, and because, overall, the material in the first chapters is the most important for people in other fields. The drawback is that arithmetic is given short shrift. This is made up for in the later chapters, which have a strong arithmetic slant. Geometry is brought back from time to time in these later chapters, in the guise of lattices, symmetry, and algebraic geometry.

Michael Artin
December 1990

A Note for the Teacher

There are few prerequisites for this book. Students should be familiar with calculus, the basic properties of the complex numbers, and mathematical induction. Some acquaintance with proofs is obviously useful, though less essential. The concepts from topology, which are used in Chapter 8, should not be regarded as prerequisites. An appendix is provided as a reference for some of these concepts; it is too brief to be suitable as a text.

Don't try to cover the book in a one-year course unless your students have already had a semester of algebra, linear algebra for instance, and are mathematically fairly mature. About a third of the material can be omitted without sacrificing much of the book's flavor, and more can be left out if necessary. The following sections, for example, would make a coherent course:

> Chapter 1, Chapter 2, Chapter 3: 1–4, Chapter 4, Chapter 5: 1–7,
> Chapter 6: 1,2, Chapter 7: 1–6, Chapter 8: 1–3,5, Chapter 10: 1–7,
> Chapter 11: 1–8, Chapter 12: 1–7, Chapter 13: 1–6.

This selection includes some of the interesting special topics: symmetry of plane figures, the geometry of SU_2, and the arithmetic of imaginary quadratic number fields. If you don't want to discuss such topics, then this is not the book for you.

It would be easy to spend an entire semester on the first four chapters, but this would defeat the purpose of the book. Since the real fun starts with Chapter 5, it is important to move along. If you plan to follow the chapters in order, try to get to that chapter as soon as is practicable, so that it can be done at a leisurely pace. It will help to keep attention focussed on the concrete examples. This is especially impor-

tant in the beginning for the students who come to the course without a clear idea of what constitutes a proof.

Chapter 1, matrix operations, isn't as exciting as some of the later ones, so it should be covered fairly quickly. I begin with it because I want to emphasize the general linear group at the start, instead of following the more customary practice of basing examples on the symmetric group. The reason for this decision is Principle 3 of the preface: The general linear group is more important.

Here are some suggestions for Chapter 2:

1. Treat the abstract material with a light touch. You can have another go at it in Chapters 5 and 6.
2. For examples, concentrate on matrix groups. Mention permutation groups only in passing. Because of their inherent notational difficulties, examples from symmetry such as the dihedral groups are best deferred to Chapter 5.
3. Don't spend too much time on arithmetic. Its natural place in this book is Chapters 10 and 11.
4. Deemphasize the quotient group construction.

Quotient groups present a pedagogical problem. While their construction is conceptually difficult, the quotient is readily presented as the image of a homomorphism in most elementary examples, and so it does not require an abstract definition. Modular arithmetic is about the only convincing example for which this is not the case. And since the integers modulo n form a ring, modular arithmetic isn't the ideal motivating example for quotients of groups. The first serious use of quotient groups comes when generators and relations are discussed in Chapter 6, and I deferred the treatment of quotients to that point in early drafts of the book. But fearing the outrage of the algebra community I ended up moving it to Chapter 2. Anyhow, if you don't plan to discuss generators and relations for groups in your course, then you can defer an in-depth treatment of quotients to Chapter 10, ring theory, where they play a central role, and where modular arithmetic becomes a prime motivating example.

In Chapter 3, vector spaces, I've tried to set up the computations with bases in such a way that the students won't have trouble keeping the indices straight. I've probably failed, but since the notation is used throughout the book, it may be advisable to adopt it.

The applications of linear operators to rotations and linear differential equations in Chapter 4 should be discussed because they are used later on, but the temptation to give differential equations their due has to be resisted. This heresy will be forgiven because you are teaching an algebra course.

There is a gradual rise in the level of sophistication which is assumed of the reader throughout the first chapters, and a jump which I've been unable to eliminate occurs in Chapter 5. Had it not been for this jump, I would have moved symmetry closer to the beginning of the book. Keep in mind that symmetry is a difficult concept. It is easy to get carried away by the material and to leave the students behind.

Except for its first two sections, Chapter 6 contains optional material. The last section on the Todd–Coxeter algorithm isn't standard; it is included to justify the discussion of generators and relations, which is pretty useless without it.

There is nothing unusual in the chapter on bilinear forms, Chapter 7. I haven't overcome the main problem with this material, that there are too many variations on the same theme, but have tried to keep the discussion short by concentrating on the real and complex cases.

In the chapter on linear groups, Chapter 8, plan to spend time on the geometry of SU_2. My students complained every year about this chapter until I expanded the sections on SU_2, after which they began asking for supplementary reading, wanting to learn more. Many of our students are not familiar with the concepts from topology when they take the course, and so these concepts require a light touch. But I've found that the problems caused by the students' lack of familiarity can be managed. Indeed, this is a good place for them to get an idea of what a manifold is. Unfortunately, I don't know a really satisfactory reference for further reading.

Chapter 9 on group representations is optional. I resisted including this topic for a number of years, on the grounds that it is too hard. But students often request it, and I kept asking myself: If the chemists can teach it, why can't we? Eventually the internal logic of the book won out and group representations went in. As a dividend, hermitian forms got an application.

The unusual topic in Chapter 11 is the arithmetic of quadratic number fields. You may find the discussion too long for a general algebra course. With this possibility in mind, I've arranged the material so that the end of Section 8, ideal factorization, is a natural stopping point.

It seems to me that one should at least mention the most important examples of fields in a beginning algebra course, so I put a discussion of function fields into Chapter 13.

There is always the question of whether or not Galois theory should be presented in an undergraduate course. It doesn't have quite the universal applicability of most of the subjects in the book. But since Galois theory is a natural culmination of the discussion of symmetry, it belongs here as an optional topic. I usually spend at least some time on Chapter 14.

I considered grading the exercises for difficulty, but found that I couldn't do it consistently. So I've only gone so far as to mark some of the harder ones with an asterisk. I believe that there are enough challenging problems, but of course one always needs more of the interesting, easier ones.

Though I've taught algebra for many years, several aspects of this book are experimental, and I would be very grateful for critical comments and suggestions from the people who use it.

"One, two, three, five, four..."
"No Daddy, it's one, two, three, four, five."
"Well if I want to say one, two, three, five, four, why can't I?"
"That's not how it goes."

Acknowledgments

Mainly, I want to thank the students who have been in my classes over the years for making them so exciting. Many of you will recognize your own contributions, and I hope that you will forgive me for not naming you individually.

Several people have used my notes in classes and made valuable suggestions—Jay Goldman, Steve Kleiman, Richard Schafer, and Joe Silverman among them. Harold Stark helped me with the number theory, and Gil Strang with the linear algebra. Also, the following people read the manuscript and commented on it: Ellen Kirkman, Al Levine, Barbara Peskin, and John Tate. I want to thank Barbara Peskin especially for reading the whole thing twice during the final year.

The figures which needed mathematical precision were made on the computer by George Fann and Bill Schelter. I could not have done them by myself.

Many thanks also to Marge Zabierek, who retyped the manuscript annually for about eight years before it was put onto the computer where I could do the revisions myself, and to Mary Roybal for her careful and expert job of editing the manuscript.

I've not consulted other books very much while writing this one, but the classics by Birkhoff and MacLane and by van der Waerden from which I learned the subject influenced me a great deal, as did Herstein's book, which I used as a text for many years. I also found some good ideas for exercises in the books by Noble and by Paley and Weichsel.

Some quotations, often out of context, are scattered about the text. I learned the Leibnitz and Russell quotes which end Chapters 5 and 6 from V. I. Arnold, and the Weyl quote which begins Chapter 8 is from Morris Klein's book *Mathematical Thought from Ancient to Modern Times*.

Chapter 1

Matrix Operations

Erſtlich wird alles dasjenige eine Größe genennt,
welches einer Vermehrung oder einer Verminderung fähig iſt,
oder wozu ſich noch etwas hinzuſetzen oder davon wegnehmen läßt.

Leonhard Euler

Matrices play a central role in this book. They form an important part of the theory, and many concrete examples are based on them. Therefore it is essential to develop facility in matrix manipulation. Since matrices pervade much of mathematics, the techniques needed here are sure to be useful elsewhere.

The concepts which require practice to handle are *matrix multiplication* and *determinants*.

1. THE BASIC OPERATIONS

Let m, n be positive integers. An $m \times n$ matrix is a collection of mn numbers arranged in a rectangular array:

$$(1.1) \qquad \overset{n \text{ columns}}{m \text{ rows} \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}}$$

For example, $\begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 5 \end{bmatrix}$ is a 2×3 matrix.

The numbers in a matrix are called the *matrix entries* and are denoted by a_{ij}, where i, j are indices (integers) with $1 \le i \le m$ and $1 \le j \le n$. The index i is called the *row index*, and j is the *column index*. So a_{ij} is the entry which appears in

the ith row and jth column of the matrix:

$$i\begin{bmatrix} \cdots & a_{ij} & \cdots\cdots \\ & \vdots & \\ & \vdots & \end{bmatrix}$$

In the example above, $a_{11} = 2$, $a_{13} = 0$, and $a_{23} = 5$.

We usually introduce a symbol such as A to denote a matrix, or we may write it as (a_{ij}).

A $1 \times n$ matrix is called an n-dimensional *row vector*. We will drop the index i when $m = 1$ and write a row vector as

(1.2) $$A = [a_1 \cdots a_n], \quad \text{or as} \quad A = (a_1, \ldots, a_n).$$

The commas in this row vector are optional. Similarly, an $m \times 1$ matrix is an m-dimensional *column vector*:

(1.3) $$B = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

A 1×1 matrix $[a]$ contains a single number, and we do not distinguish such a matrix from its entry.

(1.4) *Addition* of matrices is vector addition:

$$(a_{ij}) + (b_{ij}) = (s_{ij}),$$

where $s_{ij} = a_{ij} + b_{ij}$ for all i, j. Thus

$$\begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 5 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 3 \\ 4 & -3 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 3 \\ 5 & 0 & 6 \end{bmatrix}.$$

The sum of two matrices A, B is defined only when they are both of the same shape, that is, when they are $m \times n$ matrices with the same m and n.

(1.5) *Scalar multiplication* of a matrix by a number is defined as with vectors. The result of multiplying a number c and a matrix (a_{ij}) is another matrix:

$$c(a_{ij}) = (b_{ij}),$$

where $b_{ij} = ca_{ij}$ for all i, j. Thus

$$2\begin{bmatrix} 0 & 1 \\ 2 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 4 & 6 \\ 4 & 2 \end{bmatrix}.$$

Numbers will also be referred to as *scalars*.

The complicated notion is that of *matrix multiplication*. The first case to learn is the product AB of a row vector A (1.2) and a column vector B (1.3) which is defined when both are the same size, that is, $m = n$. Then the product AB is the 1×1 matrix or scalar

$$a_1b_1 + a_2b_2 + \cdots + a_mb_m. \tag{1.6}$$

(This product is often called the "dot product" of the two vectors.) Thus

$$\begin{bmatrix} 3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix} = 3 \cdot 1 + 1 \cdot (-1) + 2 \cdot 4 = 10.$$

The usefulness of this definition becomes apparent when we regard A and B as vectors which represent indexed quantities. For example, consider a candy bar containing m ingredients. Let a_i denote the number of grams of $(\textit{ingredient})_i$ per candy bar, and let b_i denote the cost of $(\textit{ingredient})_i$ per gram. Then the matrix product $AB = c$ computes the cost per candy bar:

$$\text{(grams/bar)} \cdot \text{(cost/gram)} = \text{(cost/bar)}.$$

On the other hand, the fact that we consider this to be the product of a row by a column is an arbitrary choice.

In general, the product of two matrices A and B is defined if the number of columns of A is equal to the number of rows of B, say if A is an $\ell \times m$ matrix and B is an $m \times n$ matrix. In this case, the product is an $\ell \times n$ matrix. Symbolically, $(\ell \times m) \cdot (m \times n) = (\ell \times n)$. The entries of the product matrix are computed by multiplying all rows of A by all columns of B, using rule (1.6) above. Thus if we denote the product AB by P, then

$$p_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj}. \tag{1.7}$$

This is the product of the ith row of A and the jth column of B.

$$i\ \begin{bmatrix} a_{i1} & \cdots & a_{im} \\ & & \end{bmatrix} \cdot \overset{j}{\begin{bmatrix} b_{1j} \\ \vdots \\ b_{mj} \end{bmatrix}} = \begin{bmatrix} & \vdots & \\ \cdots & p_{ij} & \cdots \\ & \vdots & \end{bmatrix}$$

For example,

$$\begin{bmatrix} 0 & -1 & 2 \\ 3 & 4 & -6 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}. \tag{1.8}$$

This definition of matrix multiplication has turned out to provide a very convenient computational tool.

Going back to our candy bar example, suppose that there are ℓ candy bars. Then we may form a matrix A whose ith row measures the ingredients of $(bar)_i$. If the cost is to be computed each year for n years, we may form a matrix B whose jth column measures the cost of the ingredients in $(year)_j$. The matrix product $AB = P$ computes the cost per bar: p_{ij} = cost of $(bar)_i$ in $(year)_j$.

Matrix notation was introduced in the nineteenth century to provide a shorthand way of writing linear equations. The system of equations

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad\qquad \vdots \quad &\quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

can be written in matrix notation as

$$AX = B, \tag{1.9}$$

where A denotes the coefficient matrix (a_{ij}), X and B are column vectors, and AX is the matrix product

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

Thus the matrix equation

$$\begin{bmatrix} 0 & -1 & 2 \\ 3 & 4 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

represents the following system of two equations in three unknowns:

$$\begin{aligned} -x_2 + 2x_3 &= 2 \\ 3x_1 + 4x_2 - 6x_3 &= 1. \end{aligned}$$

Equation (1.8) exhibits one solution: $x_1 = 1$, $x_2 = 4$, $x_3 = 3$.

Formula (1.7) defining the product can also be written in "sigma" notation as

$$p_{ij} = \sum_{k=1}^{m} a_{ik}b_{kj} = \sum_{k} a_{ik}b_{kj}.$$

Each of these expressions is a shorthand notation for the sum (1.7) which defines the product matrix.

Our two most important notations for handling sets of numbers are the Σ or sum notation as used above and matrix notation. The Σ notation is actually the more versatile of the two, but because matrices are much more compact we will use them whenever possible. One of our tasks in later chapters will be to translate complicated mathematical structures into matrix notation in order to be able to work with them conveniently.

Various *identities* are satisfied by the matrix operations, such as the *distributive laws*

$$(1.10) \qquad A(B + B') = AB + AB', \quad \text{and} \quad (A + A')B = AB + A'B$$

and the *associative law*

$$(1.11) \qquad (AB)C = A(BC).$$

These laws hold whenever the matrices involved have suitable sizes, so that the products are defined. For the associative law, for example, the sizes should be $A = \ell \times m$, $B = m \times n$ and, $C = n \times p$, for some ℓ, m, n, p. Since the two products (1.11) are equal, the parentheses are not required, and we will denote them by ABC. The triple product ABC is then an $\ell \times p$ matrix. For example, the two ways of computing the product

$$ABC = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$$

are

$$(AB)C = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 6 & 2 \end{bmatrix} \quad \text{and} \quad A(BC) = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 6 & 2 \end{bmatrix}.$$

Scalar multiplication is compatible with matrix multiplication in the obvious sense:

$$(1.12) \qquad c(AB) = (cA)B = A(cB).$$

The proofs of these identities are straightforward and not very interesting.

In contrast, *the commutative law does not hold* for matrix multiplication; that is,

$$(1.13) \qquad AB \neq BA, \text{ usually.}$$

In fact, if A is an $\ell \times m$ matrix and B is an $m \times \ell$ matrix, so that AB and BA are both defined, then AB is $\ell \times \ell$ while BA is $m \times m$. Even if both matrices are square, say $m \times m$, the two products tend to be different. For instance,

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \text{ while } \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Since matrix multiplication is not commutative, care must be taken when working with matrix equations. We can multiply both sides of an equation $B = C$ on the left by a matrix A, to conclude that $AB = AC$, provided that the products are defined. Similarly, if the products are defined, then we can conclude that $BA = CA$. We can not derive $AB = CA$ from $B = C$!

Any matrix all of whose entries are 0 is called a *zero matrix* and is denoted by 0, though its size is arbitrary. Maybe $0_{m \times n}$ would be better.

The entries a_{ii} of a matrix A are called its *diagonal entries,* and a matrix A is called a *diagonal matrix* if its only nonzero entries are diagonal entries.

The square $n \times n$ matrix whose only nonzero entries are 1 in each diagonal position,

(1.14)
$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 \end{bmatrix},$$

is called the $n \times n$ *identity* matrix. It behaves like 1 in multiplication: If A is an $m \times n$ matrix, then

$$I_m A = A \quad \text{and} \quad AI_n = A.$$

Here are some shorthand ways of drawing the matrix I_n:

$$I_n = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}.$$

We often indicate that a whole region in a matrix consists of zeros by leaving it blank or by putting in a single 0.

We will use $*$ to indicate an arbitrary undetermined entry of a matrix. Thus

$$\begin{bmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{bmatrix}$$

may denote a square matrix whose entries below the diagonal are 0, the other entries being undetermined. Such a matrix is called an *upper triangular* matrix.

Let A be a (square) $n \times n$ matrix. If there is a matrix B such that

(1.15)
$$AB = I_n \quad \text{and} \quad BA = I_n,$$

then B is called an *inverse* of A and is denoted by A^{-1}:

(1.16)
$$A^{-1}A = I_n = AA^{-1}.$$

When A has an inverse, it is said to be an *invertible* matrix. For example, the matrix $A = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}$ is invertible. Its inverse is $A^{-1} = \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix}$, as is seen by computing

the products AA^{-1} and $A^{-1}A$. Two more examples are:

$$\begin{bmatrix} 1 & \\ & 2 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & \\ & \frac{1}{2} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -1 \\ & 1 \end{bmatrix}.$$

We will see later that A is invertible if there is a matrix B such that either one of the two relations $AB = I_n$ or $BA = I_n$ holds, and that B is then the inverse [see (2.23)]. But since multiplication of matrices is not commutative, this fact is not obvious. It fails for matrices which aren't square. For example, let $A = [1 \quad 2]$ and let $B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Then $AB = [1] = I_1$, but $BA = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \neq I_2$.

On the other hand, an inverse is unique if it exists at all. In other words, there can be only one inverse. Let B, B' be two matrices satisfying (1.15), for the same matrix A. We need only know that $AB = I_n$ (B is a *right inverse*) and that $B'A = I_n$ (B' is a *left inverse*). By the associative law, $B'(AB) = (B'A)B$. Thus

$$(1.17) \qquad B' = B'I = B'(AB) = (B'A)B = IB = B,$$

and so $B' = B$. □

(1.18) **Proposition.** Let A, B be $n \times n$ matrices. If both are invertible, so is their product AB, and

$$(AB)^{-1} = B^{-1}A^{-1}.$$

More generally, if $A_1, \ldots, A_m$ are invertible, then so is the product $A_1 \cdots A_m$, and its inverse is $A_m^{-1} \cdots A_1^{-1}$.

Thus the inverse of $\begin{bmatrix} 1 & \\ & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ & 2 \end{bmatrix}$ is $\begin{bmatrix} 1 & -1 \\ & 1 \end{bmatrix}\begin{bmatrix} 1 & \\ & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & -\frac{1}{2} \\ & \frac{1}{2} \end{bmatrix}$.

Proof. Assume that A, B are invertible. Then we check that $B^{-1}A^{-1}$ is the inverse of AB:

$$ABB^{-1}A^{-1} = AIA^{-1} = AA^{-1} = I,$$

and similarly

$$B^{-1}A^{-1}AB = \cdots = I.$$

The last assertion is proved by induction on m [see Appendix (2.3)]. When $m = 1$, the assertion is that if A_1 is invertible then A_1^{-1} is the inverse of A_1, which is trivial. Next we assume that the assertion is true for $m = k$, and we proceed to check it for $m = k + 1$. We suppose that $A_1, \ldots, A_{k+1}$ are invertible $n \times n$ matrices, and we denote by P the product $A_1 \cdots A_k$ of the first k matrices. By the induction hypothesis, P is invertible, and its inverse is $A_k^{-1} \cdots A_1^{-1}$. Also, A_{k+1} is invertible. So, by what has been shown for two invertible matrices, the product $PA_{k+1} = A_1 \cdots A_k A_{k+1}$ is invertible, and its inverse is $A_{k+1}^{-1}P^{-1} = A_{k+1}^{-1}A_k^{-1} \cdots A_1^{-1}$. This shows that the assertion is true for $m = k + 1$, which completes the induction proof. □

Though this isn't clear from the definition of matrix multiplication, we will see that most square matrices are invertible. But finding the inverse explicitly is not a simple problem when the matrix is large.

The set of all invertible $n \times n$ matrices is called the n-dimensional *general linear group* and is denoted by GL_n. The general linear groups will be among our most important examples when we study the basic concept of a group in the next chapter.

Various tricks simplify matrix multiplication in favorable cases. *Block multiplication* is one of them. Let M, M' be $m \times n$ and $n \times p$ matrices, and let r be an integer less than n. We may decompose the two matrices into blocks as follows:

$$M = [A \,|\, B] \quad \text{and} \quad M' = \left[\begin{array}{c} A' \\ \hline B' \end{array}\right],$$

where A has r columns and A' has r rows. Then the matrix product can be computed as follows:

$$MM' = AA' + BB'. \tag{1.19}$$

This decomposition of the product follows directly from the definition of multiplication, and it may facilitate computation. For example,

$$\left[\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & 7 \end{array}\right]\left[\begin{array}{cc} 2 & 3 \\ 4 & 8 \\ \hline 0 & 0 \end{array}\right] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 3 \\ 4 & 8 \end{bmatrix} + \begin{bmatrix} 5 \\ 7 \end{bmatrix}[0 \quad 0] = \begin{bmatrix} 2 & 3 \\ 4 & 8 \end{bmatrix}.$$

Note that formula (1.19) looks the same as rule (1.6) for multiplying a row vector and a column vector.

We may also multiply matrices divided into more blocks. For our purposes, a decomposition into four blocks will be the most useful. In this case the rule for block multiplication is the same as for multiplication of 2×2 matrices. Let $r + s = n$ and let $k + \ell = m$. Suppose we decompose an $m \times n$ matrix M and an $n \times p$ matrix M' into submatrices

$$M = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right], \quad M' = \left[\begin{array}{c|c} A' & B' \\ \hline C' & D' \end{array}\right],$$

where the number of columns of A is equal to the number of rows of A'. Then the rule for block multiplication is

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]\left[\begin{array}{c|c} A' & B' \\ \hline C' & D' \end{array}\right] = \left[\begin{array}{c|c} AA' + BC' & AB' + BD' \\ \hline CA' + DC' & CB' + DD' \end{array}\right]. \tag{1.20}$$

For example,

$$\left[\begin{array}{cc|c} 1 & 0 & 5 \\ \hline 0 & 1 & 7 \end{array}\right] \cdot \left[\begin{array}{cc|cc} 2 & 3 & 1 & 1 \\ 4 & 1 & 0 & 0 \\ \hline 0 & 1 & 1 & 0 \end{array}\right] = \left[\begin{array}{cc|cc} 2 & 8 & 6 & 1 \\ \hline 4 & 8 & 7 & 0 \end{array}\right].$$

In this product, the upper left block is $[1 \quad 0]\begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} + [5][0 \quad 1] = [2 \quad 8]$, etc.

Again, this rule can be verified directly from the definition of matrix multiplication. In general, block multiplication can be used whenever two matrices are decomposed into submatrices in such a way that the necessary products are defined.

Besides facilitating computations, block multiplication is a useful tool for proving facts about matrices by induction.

2. ROW REDUCTION

Let $A = (a_{ij})$ be an $m \times n$ matrix, and consider a variable $n \times p$ matrix $X = (x_{ij})$. Then the matrix equation

(2.1) $$Y = AX$$

defines the $m \times p$ matrix $Y = (y_{ij})$ as a function of X. This operation is called *left multiplication by A*:

(2.2) $$y_{ij} = a_{i1}x_{1j} + \cdots + a_{in}x_{nj}.$$

Notice that in formula (2.2) the entry y_{ij} depends only on $x_{1j}, \ldots, x_{nj}$, that is, on the jth column of X and on the ith row of the matrix A. Thus A operates separately on each column of X, and we can understand the way A operates by considering its action on column vectors:

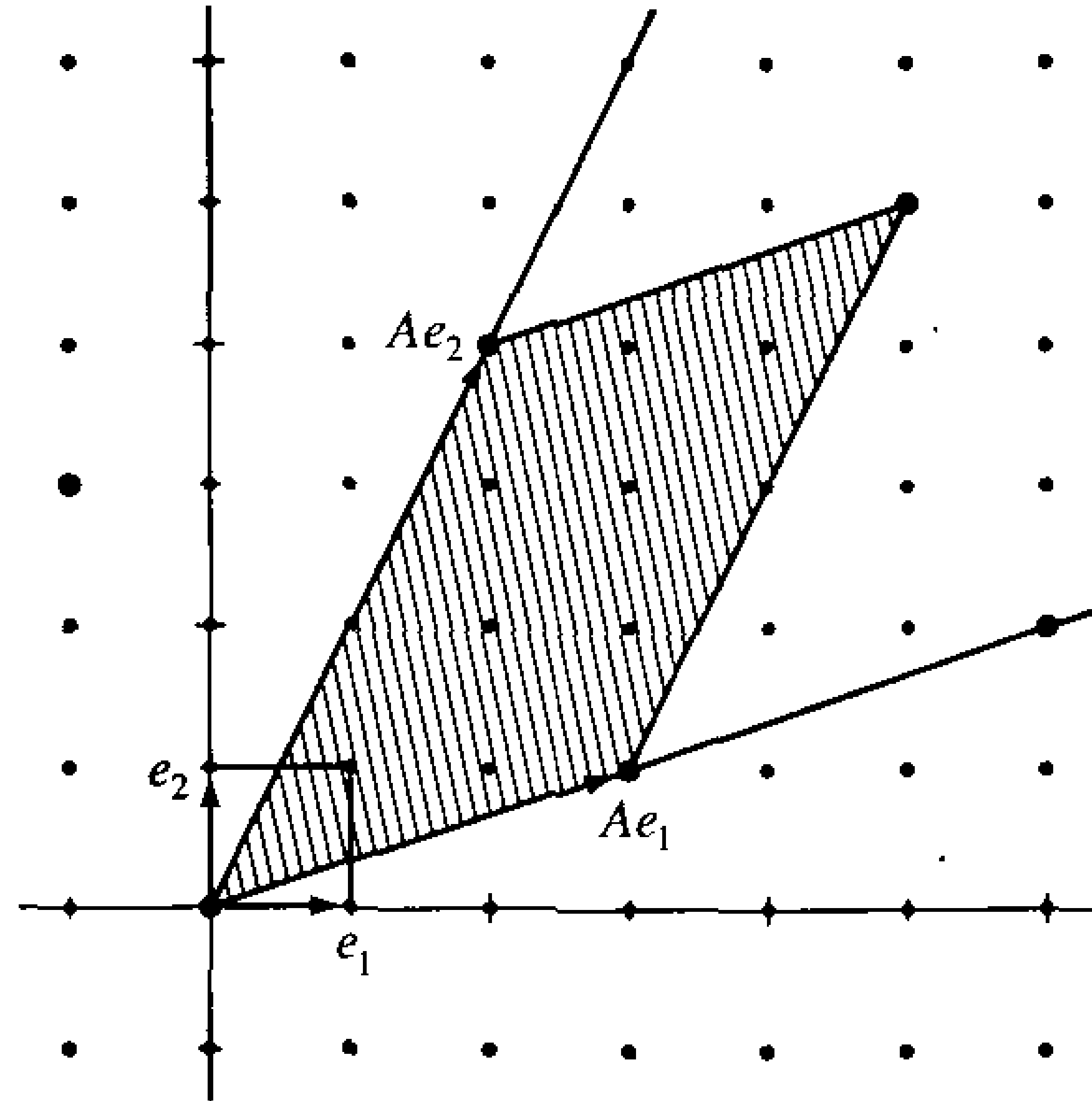
$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}.$$

Left multiplication by A on column vectors can be thought of as a function from the space of n-dimensional column vectors X to the space of m-dimensional column vectors Y, or a collection of m functions of n variables:

$$y_i = a_{i1}x_1 + \cdots + a_{in}x_n \qquad (i = 1, \ldots, m).$$

It is called a *linear transformation,* because the functions are homogeneous and linear. (A *linear* function of a set of variables $u_1, \ldots, u_k$ is one of the form $a_1u_1 + \cdots + a_ku_k + c$, where $a_1, \ldots, a_k, c$ are scalars. Such a function is *homogeneous linear* if the constant term c is zero.)

A picture of the operation of the 2×2 matrix $\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$ is shown below. It maps 2-space to 2-space:



(2.3) **Figure.**

Going back to the operation of A on an $n \times p$ matrix X, we can interpret the fact that A acts in the same way on each column of X as follows: Let Y_i denote the ith row of Y, which we view as a *row vector:*

$$Y = \begin{bmatrix} \text{---} & Y_1 & \text{---} \\ \text{---} & Y_2 & \text{---} \\ & \vdots & \\ \text{---} & Y_p & \text{---} \end{bmatrix}$$

We can compute Y_i in terms of the rows X_j of X, in vector notation, as

(2.4) $$Y_i = a_{i1}X_1 + \cdots + a_{in}X_n.$$

This is just a restatement of (2.2), and it is another example of block multiplication. For example, the bottom row of the product

$$\begin{bmatrix} 0 & -1 & 2 \\ 3 & 4 & -6 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 4 & 2 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 1 & -4 \end{bmatrix}$$

can be computed as $3[1 \quad 0] + 4[4 \quad 2] - 6[3 \quad 2] = [1 \quad -4]$.

When A is a square matrix, we often speak of left multiplication by A as a *row operation*.

The simplest nonzero matrices are the matrix units, which we denote by e_{ij}:

(2.5) $$e_{ij} = i \begin{bmatrix} & \overset{j}{\cdot} & & & \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ & \vdots & & & \end{bmatrix}.$$

This matrix e_{ij} has a 1 in the (i, j) position as its only nonzero entry. (We usually denote matrices by capital letters, but the use of a small letter for the matrix units is traditional.) Matrix units are useful because every matrix $A = (a_{ij})$ can be written out as a sum in the following way:

$$A = a_{11}e_{11} + a_{12}e_{12} + \cdots + a_{nn}e_{nn} = \sum_{i,j} a_{ij}e_{ij}.$$

The indices i, j under the sigma mean that the sum is to be taken over all values of i and all values of j. For instance

$$\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 3 & \\ & \end{bmatrix} + \begin{bmatrix} & 2 \\ & \end{bmatrix} + \begin{bmatrix} & \\ 1 & \end{bmatrix} + \begin{bmatrix} & \\ & 4 \end{bmatrix} = 3e_{11} + 2e_{12} + 1e_{21} + 4e_{22}.$$

Such a sum is called a linear combination of the matrices e_{ij}.

The matrix units are convenient for the study of addition and scalar multiplication of matrices. But to study matrix multiplication, some square matrices called *elementary matrices* are more useful. There are three types of elementary matrix:

(2.6i)
$$\begin{bmatrix} 1 & & & \\ & \cdot & & a \\ & & \ddots & \\ & & & 1 \end{bmatrix} \text{ or } \begin{bmatrix} 1 & & & \\ & \cdot & & \\ & & \ddots & \\ & a & & 1 \end{bmatrix} = I + ae_{ij} \qquad (i \neq j).$$

Such a matrix has diagonal entries 1 and one nonzero off-diagonal entry.

(2.6ii)
$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 0 & & 1 & \\ & & & \ddots & & \\ & & 1 & & 0 & \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix} = I + e_{ij} + e_{ji} - e_{ii} - e_{jj}.$$

Here the ith and jth diagonal entries of I are replaced by zero, and two 1's are added in the (i, j) and (j, i) positions. (The formula in terms of the matrix units is rather ugly, and we won't use it much.)

(2.6iii)
$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & c & & \\ & & & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix} = I + (c - 1)e_{ii}, \qquad (c \neq 0).$$

One diagonal entry of the identity matrix is replaced by a nonzero number c.

The elementary 2×2 matrices are

$$\text{(i)} \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix}, \quad \text{(ii)} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \text{(iii)} \begin{bmatrix} c & \\ & 1 \end{bmatrix}, \begin{bmatrix} 1 & \\ & c \end{bmatrix},$$

where, as above, a is arbitrary and c is an arbitrary nonzero number.

The elementary matrices E operate on a matrix X as described below.

(2.7) To get the matrix EX, you must:

Type (i): *Replace the* ith *row* X_i *by* $X_i + aX_j$, or *add* $a \cdot$(row j) *to* (row i);
Type (ii): *Interchange* (row i) *and* (row j);
Type (iii): *Multiply* (row i) *by a nonzero scalar* c.

These operations are called *elementary row operations*. Thus multiplication by an elementary matrix is an elementary row operation. You should verify these rules of multiplication carefully.

(2.8) **Lemma.** Elementary matrices are invertible, and their inverses are also elementary matrices.

The proof of this lemma is just a calculation. The inverse of an elementary matrix is the matrix corresponding to the inverse row operation: If $E = I + ae_{ij}$ is of Type (i), then $E^{-1} = I - ae_{ij}$; "*subtract* $a \cdot$(row j) *from* (row i)". If E is of Type (ii), then $E^{-1} = E$, and if E is of Type (iii), then E^{-1} is of the same type, with c^{-1} in the position that c has in E; "*multiply* (row i) *by* c^{-1}". □

We will now study the effect of elementary row operations (2.7) on a matrix A, with the aim of ending up with a simpler matrix A':

$$A \xrightarrow{\text{sequence of operations}} \cdots \longrightarrow A'.$$

Since each elementary row operation is obtained as the result of multiplication by an elementary matrix, we can express the result of a succession of such operations as multiplication by a sequence $E_1, \ldots, E_k$ of elementary matrices:

(2.9)
$$A' = E_k \cdots E_2 E_1 A.$$

This procedure is called *row reduction,* or *Gaussian elimination*. For example, we can simplify the matrix

(2.10)
$$M = \begin{bmatrix} 1 & 0 & 2 & 1 & 5 \\ 1 & 1 & 5 & 2 & 7 \\ 1 & 2 & 8 & 4 & 12 \end{bmatrix}$$

by using the first type of elementary operation to clear out as many entries as possible:

$$\begin{bmatrix} 1 & 0 & 2 & 1 & 5 \\ 1 & 1 & 5 & 2 & 7 \\ 1 & 2 & 8 & 4 & 12 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 2 & 1 & 5 \\ 0 & 1 & 3 & 1 & 2 \\ 1 & 2 & 8 & 4 & 12 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 2 & 1 & 5 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 2 & 6 & 3 & 7 \end{bmatrix} \longrightarrow$$

$$\begin{bmatrix} 1 & 0 & 2 & 1 & 5 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{bmatrix} \longrightarrow \longrightarrow \begin{bmatrix} 1 & 0 & 2 & 0 & 2 \\ 0 & 1 & 3 & 0 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{bmatrix}.$$

Row reduction is a useful method of solving systems of linear equations. Suppose we are given a system of m equations in n unknowns, say $AX = B$ as in (1.9), where A is an $m \times n$ matrix, X is an unknown column vector, and B is a given column vector. To solve this system, we form the $m \times (n + 1)$ block matrix

$$M = [A \mid B] = \left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_n \end{array}\right], \tag{2.11}$$

and we perform row operations to simplify M. Note that $EM = [EA \mid EB]$. Let

$$M' = [A' \mid B']$$

be the result of a sequence of row operations. The key observation follows:

(2.12) **Proposition.** The solutions of $A'X = B'$ are the same as those of $AX = B$.

Proof. Since M' is obtained by a sequence of elementary row operations,

$$M' = E_r \cdots E_1 M.$$

Let $P = E_r \cdots E_1$. This matrix is invertible, by Lemma (2.8) and Proposition (1.18). Also, $M' = [A' \mid B'] = [PA \mid PB]$. If X is a solution of the original system $AX = B$, then $PAX = PB$, which is to say, $A'X = B'$. So X also solves the new system. Conversely, if $A'X = B'$, then $AX = P^{-1}A'X = P^{-1}B' = B$, so X solves the system $AX = B$ too. □

For example, consider the system

$$\begin{aligned} x_1 + \qquad 2x_3 + \ x_4 &= 5 \\ x_1 + \ x_2 + 5x_3 + 2x_4 &= 7 \\ x_1 + 2x_2 + 8x_3 + 4x_4 &= 12. \end{aligned} \tag{2.13}$$

Its augmented matrix is the matrix M considered above (2.10), so our row reduction of this matrix shows that this system of equations is equivalent to

$$\begin{aligned} x_1 + \quad 2x_3 \qquad &= 2 \\ x_2 + 3x_3 \qquad &= -1 \\ x_4 &= 3. \end{aligned}$$

We can read off the solutions of this system immediately: We may choose x_3 arbitrarily and then solve for x_1, x_2, and x_4. The general solution of (2.13) can therefore be written in the form

$$x_3 = c_3,\ x_1 = 1 - 2c_3,\ x_2 = -1 - 3c_3,\ x_4 = 3,$$

where c_3 is arbitrary.

We now go back to row reduction of an arbitrary matrix. It is not hard to see that, by a sequence of row operations, any matrix A can be reduced to one which looks roughly like this:

$$\text{(2.14)} \qquad A = \left[\begin{array}{cccccccc} 1 & *..* & 0 & *..* & 0 & *...* & 0 & \\ & & 1 & *..* & 0 & *...* & 0 & \\ & & & & 1 & *...* & 0 & \cdots \\ & & & & & & 1 & \\ & & & & & & & \ddots \end{array}\right],$$

where $*$ denotes an arbitrary number and the large blank space consists of zeros. This is called a *row echelon matrix*. For instance,

$$\begin{bmatrix} 1 & 6 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

is a row echelon matrix. So is the end result of our reduction of (2.10). The definition of a row echelon matrix is given in (2.15):

(2.15)

(a) The first nonzero entry in every row is 1. This entry is called a *pivot*.
(b) The first nonzero entry of row $i + 1$ is to the right of the first nonzero entry of row i.
(c) The entries above a pivot are zero.

To make a row reduction, find the first column which contains a nonzero entry. (If there is none, then $A = 0$, and 0 is a row echelon matrix.) Interchange rows using an elementary operation of Type (ii), moving a nonzero entry to the top row. Normalize this entry to 1 using an operation of Type (iii). Then clear out the other entries in its column by a sequence of operations of Type (i). The resulting matrix will have the block form

$$\left[\begin{array}{ccc|c|ccc} 0 & \dots & 0 & 1 & * & \dots\dots & * \\ \hline 0 & \dots & 0 & 0 & * & \dots\dots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & * & \dots\dots & * \end{array}\right], \text{ which we may write as } \left[\begin{array}{c|c|c} & 1 & B \\ \hline & & D \end{array}\right] = A'.$$

We now continue, performing row operations on the smaller matrix D (cooking until done). Formally, this is induction on the size of the matrix. The principle of complete induction [see Appendix (2.6)] allows us to assume that every matrix with fewer rows than A can be reduced to row echelon form. Since D has fewer rows, we may assume that it can be reduced to a row echelon matrix, say D''. The row operations we perform to reduce D to D'' will not change the other blocks making up A'. Therefore A' can be reduced to the matrix

$$\left[\begin{array}{c|c|c} & 1 & B \\ \hline & & D'' \end{array}\right] = A'',$$

which satisfies requirements (2.15a and b) for a row echelon matrix. Therefore our original matrix A can be reduced to this form. The entries in B above the pivots of D'' can be cleared out at this time, to finish the reduction to row echelon form. □

It can be shown that the row echelon matrix obtained from a given matrix A by row reduction is unique, that is, that it does not depend on the particular sequence of operations used. However, this is not a very important point, so we omit the proof.

The reason that row reduction is useful is that we can solve a system of equations $A'X = B'$ immediately if A' is in row echelon form. For example, suppose that

$$[A'|B'] = \left[\begin{array}{cccc|c} 1 & 6 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

There is no solution to $A'X = B'$ because the third equation is $0 = 1$. On the other hand,

$$[A'|B'] = \left[\begin{array}{cccc|c} 1 & 6 & 0 & 1 & 1 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

has solutions. Choosing x_2, x_4 arbitrarily, we can solve the first equation for x_1 and the second for x_3. This is the procedure we use to solve system (2.13).

The general rule is as follows:

(2.16) **Proposition.** Let $M' = [A'|B']$ be a row echelon matrix. Then the system of equations $A'X = B'$ has a solution if and only if there is no pivot in the last column B'. In that case, an arbitrary value can be assigned to the unknown x_i if column i does not contain a pivot. □

Of course every *homogeneous* linear system $AX = 0$ has the trivial solution $X = 0$. But looking at the row echelon form again, we can conclude that if there are more unknowns than equations then the homogeneous equation $AX = 0$ has a *nontrivial* solution for X :

(2.17) **Corollary.** Every system $AX = 0$ of m homogeneous equations in n unknowns, with $m < n$, has a solution X in which some x_i is nonzero.

For, let $A'X = 0$ be the associated row echelon equation, and let r be the number of pivots of A'. Then $r \leq m$. According to the proposition, we may assign arbitrary values to $n - r$ variables x_i. □

We will now use row reduction to characterize square invertible matrices.

(2.18) **Proposition.** Let A be a square matrix. The following conditions are equivalent:

(a) A can be reduced to the identity by a sequence of elementary row operations.
(b) A is a product of elementary matrices.
(c) A is invertible.
(d) The system of homogeneous equations $AX = 0$ has only the trivial solution $X = 0$.

Proof. We will prove this proposition by proving the implications (a)⇒(b)⇒(c)⇒(d)⇒(a). To show that (a) implies (b), suppose that A can be reduced to the identity by row operations: $E_k \cdots E_1 A = I$. Multiplying both sides of this equation on the left by $E_1^{-1} \cdots E_k^{-1}$, we obtain $A = E_1^{-1} \cdots E_k^{-1}$. Since the inverse of an elementary matrix is elementary, this shows that A is a product of elementary matrices. Because a product of elementary matrices is invertible, (b) implies (c). If A is invertible we can multiply both sides of the equation $AX = 0$ by A^{-1} to derive $X = 0$. So the equation $AX = 0$ has only the trivial solution. This shows that (c) implies (d).

To prove the last implication, that (d) implies (a), we take a look at *square* row echelon matrices M. We note the following dichotomy:

(2.19) *Let M be a square row echelon matrix.*
Either M is the identity matrix, or its bottom row is zero.

This is easy to see, from (2.15).

Suppose that (a) does not hold for a given matrix A. Then A can be reduced by row operations to a matrix A' whose bottom row is zero. In this case there are at most $n-1$ nontrivial equations in the linear system $A'X = 0$, and so Corollary (2.17) tells us that this system has a nontrivial solution. Since the equation $AX = 0$ is equivalent to $A'X = 0$, it has a nontrivial solution as well. This shows that if (a) fails then (d) does too; hence (d) implies (a). This completes the proof of Proposition (2.18). □

(2.20) **Corollary.** If a row of a square matrix A is zero, then A is not invertible. □

Row reduction provides a method of computing the inverse of an invertible matrix A: We reduce A to the identity by row operations:

$$E_k \cdots E_1 A = I$$

as above. Multiplying both sides of this equation on the right by A^{-1}, we have

$$E_k \cdots E_1 I = A^{-1}.$$

(2.21) **Corollary.** Let A be an invertible matrix. To compute its inverse A^{-1}, apply elementary row operations $E_1, \ldots, E_k$ to A, reducing it to the identity matrix. The same sequence of operations, when applied to I, yields A^{-1}.

The corollary is just a restatement of the two equations. □

(2.22) **Example.** We seek the inverse of the matrix

$$A = \begin{bmatrix} 5 & 4 \\ 6 & 5 \end{bmatrix}.$$

To compute it we form the 2×4 block matrix

$$[A \mid I] = \left[\begin{array}{cc|cc} 5 & 4 & 1 & 0 \\ 6 & 5 & 0 & 1 \end{array}\right].$$

We perform row operations to reduce A to the identity, carrying the right side along, and thereby end up with A^{-1} on the right because of Corollary (2.21).

$$[A \mid I] = \left[\begin{array}{cc|cc} 5 & 4 & 1 & 0 \\ 6 & 5 & 0 & 1 \end{array}\right] \qquad \text{Subtract (row 1) from (row 2)}$$

$$\longrightarrow \left[\begin{array}{cc|cc} 5 & 4 & 1 & 0 \\ 1 & 1 & -1 & 1 \end{array}\right] \qquad \text{Subtract } 4 \cdot \text{(row 2) from (row 1)}$$

$$\longrightarrow \left[\begin{array}{cc|cc} 1 & 0 & 5 & -4 \\ 1 & 1 & -1 & 1 \end{array}\right] \qquad \text{Subtract (row 1) from (row 2)}$$

$$\longrightarrow \left[\begin{array}{cc|cc} 1 & 0 & 5 & -4 \\ 0 & 1 & -6 & 5 \end{array}\right] = [I \mid A^{-1}].$$

Thus $A^{-1} = \begin{bmatrix} 5 & -4 \\ -6 & 5 \end{bmatrix}$.

(2.23) **Proposition.** Let A be a square matrix which has either a left inverse B: $BA = I$, or a right inverse: $AB = I$. Then A is invertible, and B is its inverse.

Proof. Suppose that $AB = I$. We perform row reduction on A. According to (2.19), there are elementary matrices $E_1, \ldots, E_k$ so that $A' = E_k \ldots E_1 A$ either is the

identity matrix or has bottom row zero. Then $A'B = E_k \dots E_1$, which is an invertible matrix. Hence the bottom row of $A'B$ is not zero, and it follows that A' has a nonzero bottom row too. So $A' = I$. By (2.18), A is invertible, and the equations $I = E_k \dots E_1 A$ and $AB = I$ show that $A^{-1} = E_k \dots E_1 = B$ (see (1.17)). The other case is that $BA = I$. Then we can interchange A and B in the above argument and conclude that B is invertible and A is its inverse. So A is invertible too. □

For most of this discussion, we could have worked with columns rather than rows. We chose to work with rows in order to apply the results to systems of linear equations; otherwise columns would have served just as well. Rows and columns are interchanged by the matrix *transpose*. The transpose of an $m \times n$ matrix A is the $n \times m$ matrix A^t obtained by reflecting about the diagonal: $A^t = (b_{ij})$, where

$$b_{ij} = a_{ji}.$$

For instance,

$$\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}^t = \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix} \quad \text{and} \quad [1 \quad 2 \quad 3]^t = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

The rules for computing with the transpose are given in (2.24):

(2.24)

(a) $$(A + B)^t = A^t + B^t.$$

(b) $$(cA)^t = cA^t.$$

(c) $$(AB)^t = B^t A^t!$$

(d) $$(A^t)^t = A.$$

Using formulas (2.24c and d), we can deduce facts about *right multiplication*, XP, from the corresponding facts about left multiplication.

The elementary matrices (2.6) act by right multiplication as the following *elementary column operations:*

(2.25)

(a) *Add* $a \cdot$ (column i) *to* (column j).
(b) *Interchange* (column i) *and* (column j).
(c) *Multiply* (column i) *by* $c \neq 0$.

3. DETERMINANTS

Every square matrix A has a number associated to it called its *determinant*. In this section we will define the determinant and derive some of its properties. The determinant of a matrix A will be denoted by $\det A$.

The determinant of a 1×1 matrix is just its unique entry

$$\det [a] = a, \tag{3.1}$$

and the determinant of a 2×2 matrix is given by the formula

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc. \tag{3.2}$$

If we think of a 2×2 matrix A as an operator on the space $\mathbb{R}^2$ of real two-dimensional vectors, as in Section 2, then $\det A$ can be interpreted geometrically. Its absolute value is the area of the parallelogram which forms the image of a unit square under the operation. For example, the area of the shaded region of Figure (2.3) is 10. The determinant is positive or negative according to whether the orientation of the square is preserved or reversed by the operation. Moreover, $\det A = 0$ if and only if the parallelogram degenerates to a line segment, and this occurs if and only if the two columns of A are proportional.

The set of all $n \times n$ matrices forms a space of dimension n^2, which we denote by $\mathbb{R}^{n \times n}$. We will regard the determinant of $n \times n$ matrices as a *function* from this space to the real numbers:

$$\det\colon \mathbb{R}^{n \times n} \longrightarrow \mathbb{R}.$$

This just means that det is a function of the n^2 matrix entries. There is one such function for each positive integer n. Unfortunately there are many formulas for the determinant, and all of them are complicated when n is large. The determinant is important because it has very nice properties, though there is no simple formula for it. Not only are the formulas complicated, but it may not be easy to show directly that two of them define the same function. So we will use the following strategy: We choose one formula essentially at random and take it as the definition of the determinant. In that way we are talking about a particular function. We show that the function we have chosen has certain very special properties. We also show that our chosen function is the *only* one having these properties. Then, to check that some other formula defines the same determinant, we have to check only that the function which it defines has these same properties. It turns out that this is usually relatively easy.

The determinant of an $n \times n$ matrix can be computed in terms of certain $(n-1) \times (n-1)$ determinants by a process called *expansion by minors*. This expansion allows us to give a recursive definition of the determinant function. Let A be an $n \times n$ matrix and let A_{ij} denote the $(n-1) \times (n-1)$ matrix obtained by crossing out the ith row and the jth column of A:

(3.3) [diagram: square with row i and column j hatched out] $= A_{ij}.$

For example, if

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 2 \\ 0 & 5 & 1 \end{bmatrix}, \text{ then } \quad A_{21} = \begin{bmatrix} 0 & 3 \\ 5 & 1 \end{bmatrix}.$$

Expansion by minors on the first column is the formula

(3.4) $$\det A = a_{11} \det A_{11} - a_{21} \det A_{21} +, - \cdots \pm a_{n1} \det A_{n1}.$$

The signs alternate. We take this formula, together with (3.1), as a recursive *definition of the determinant*. Notice that the formula agrees with (3.2) for 2×2 matrices.

The determinant of the matrix A shown above is

$$\det A = 1 \cdot \det\begin{bmatrix} 1 & 2 \\ 5 & 1 \end{bmatrix} - 2 \cdot \det\begin{bmatrix} 0 & 3 \\ 5 & 1 \end{bmatrix} + 0 \cdot \det\begin{bmatrix} 0 & 3 \\ 1 & 2 \end{bmatrix}.$$

The three 2×2 determinants which appear here can be computed by expanding by minors again and using (3.1), or by using (3.2), to get

$$\det A = 1 \cdot (-9) - 2 \cdot (-15) + 0 \cdot (-3) = 21.$$

There are other formulas for the determinant, including expansions by minors on other columns and on rows, which we will derive presently [see (4.11, 5.1, 5.2)].

It is important, both for computation of determinants and for theoretical considerations, to know some of the many special properties satisfied by determinants. Most of them can be verified by direct computation and induction on n, using expansion by minors (3.4). We will list some without giving formal proofs. In order to be able to interpret these properties for functions other than the determinant, we will denote the determinant by the symbol d for the time being.

(3.5) $$\mathrm{d}(I) = 1.$$

(3.6) *The function* $\mathrm{d}(A)$ *is linear in the rows of the matrix.*

By this we mean the following: Let R_i denote the row vector which is the ith row of the matrix, so that A can be written symbolically as

$$A = \begin{bmatrix} \text{---}\, R_1 \,\text{---} \\ \vdots \\ \text{---}\, R_n \,\text{---} \end{bmatrix}.$$

By definition, *linearity* in the ith row means that whenever R and S are row vectors then

$$\mathrm{d}\begin{bmatrix} \vdots \\ \text{---}\, R+S \,\text{---} \\ \vdots \end{bmatrix} = \mathrm{d}\begin{bmatrix} \vdots \\ \text{---}\, R \,\text{---} \\ \vdots \end{bmatrix} + \mathrm{d}\begin{bmatrix} \vdots \\ \text{---}\, S \,\text{---} \\ \vdots \end{bmatrix},$$

and

$$\mathrm{d}\begin{bmatrix} \vdots \\ \text{—} cR \text{—} \\ \vdots \end{bmatrix} = c\,\mathrm{d}\begin{bmatrix} \vdots \\ \text{—} R \text{—} \\ \vdots \end{bmatrix},$$

where the other rows of the matrices appearing in these relations are the same throughout. For example,

$$\det\begin{bmatrix} 1 & 2 & 4 \\ 3+5 & 4+6 & 2+3 \\ 2 & -1 & 0 \end{bmatrix} = \det\begin{bmatrix} 1 & 2 & 4 \\ 3 & 4 & 2 \\ 2 & -1 & 0 \end{bmatrix} + \det\begin{bmatrix} 1 & 2 & 4 \\ 5 & 6 & 3 \\ 2 & -1 & 0 \end{bmatrix},$$

and

$$\det\begin{bmatrix} 1 & 2 & 4 \\ 2\cdot 5 & 2\cdot 6 & 2\cdot 3 \\ 2 & -1 & 0 \end{bmatrix} = 2 \cdot \det\begin{bmatrix} 1 & 2 & 4 \\ 5 & 6 & 3 \\ 2 & -1 & 0 \end{bmatrix}.$$

Linearity allows us to operate on *one row at a time,* with the other rows left fixed.

Another property:

(3.7) *If two adjacent rows of a matrix A are equal, then* $\mathrm{d}(A) = 0$.

Let us prove this fact by induction on n. Suppose that rows j and $j + 1$ are equal. Then the matrices A_{i1} defined by (3.3) also have two rows equal, except when $i = j$ or $i = j + 1$. When A_{i1} has two equal rows, its determinant is zero by induction. Thus only two terms of (3.4) are different from zero, and

$$\mathrm{d}(A) = \pm a_{j1}\mathrm{d}(A_{j1}) \mp a_{j+1\,1}\mathrm{d}(A_{j+1\,1}).$$

Moreover, since the rows R_j and R_{j+1} are equal, it follows that $A_{j1} = A_{j+1\,1}$ and that $a_{j1} = a_{j+1\,1}$. Since the signs alternate, the two terms on the right side cancel, and the determinant is zero.

Properties (3.5–3.7) characterize determinants uniquely [see (3.14)], and we will derive further relations from them without going back to definition (3.4).

(3.8) *If a multiple of one row is added to an adjacent row, the determinant is unchanged.*

For, by (3.6) and (3.7),

$$\mathrm{d}\begin{bmatrix} \vdots \\ \text{—} R \text{—} \\ \text{—} S+cR \text{—} \\ \vdots \end{bmatrix} = \mathrm{d}\begin{bmatrix} \vdots \\ \text{—} R \text{—} \\ \text{—} S \text{—} \\ \vdots \end{bmatrix} + c\,\mathrm{d}\begin{bmatrix} \vdots \\ \text{—} R \text{—} \\ \text{—} R \text{—} \\ \vdots \end{bmatrix} = \mathrm{d}\begin{bmatrix} \vdots \\ \text{—} R \text{—} \\ \text{—} S \text{—} \\ \vdots \end{bmatrix}.$$

The same reasoning works if S is above R.

(3.9) *If two adjacent rows are interchanged, the determinant is multiplied by* -1.

We apply (3.8) repeatedly:

$$d\begin{bmatrix} \vdots \\ \text{---}R\text{---} \\ \text{---}S\text{---} \\ \vdots \end{bmatrix} = d\begin{bmatrix} \vdots \\ \text{---}R\text{---} \\ \text{---}(S-R)\text{---} \\ \vdots \end{bmatrix} = d\begin{bmatrix} \vdots \\ \text{---}R+(S-R)\text{---} \\ \text{---}(S-R)\text{---} \\ \vdots \end{bmatrix}$$

$$= d\begin{bmatrix} \vdots \\ \text{---}S\text{---} \\ \text{---}(S-R)\text{---} \\ \vdots \end{bmatrix} = d\begin{bmatrix} \vdots \\ \text{---}S\text{---} \\ \text{---}(-R)\text{---} \\ \vdots \end{bmatrix} = -d\begin{bmatrix} \vdots \\ \text{---}S\text{---} \\ \text{---}R\text{---} \\ \vdots \end{bmatrix}.$$

(3.7′) *If two rows of a matrix A are equal, then* $d(A) = 0$.

For, interchanging adjacent rows a few times results in a matrix A' with two adjacent rows equal. By (3.7) $d(A') = 0$, and by (3.9) $d(A) = \pm\det(A')$.

Using (3.7′), the proofs of (3.8) and (3.9) show the following:

(3.8′) *If a multiple of one row is added to another row, the determinant is not changed.*

(3.9′) *If two rows are interchanged, the determinant is multiplied by* -1.

Also, (3.6) implies the following:

(3.10) *If a row of A is zero, then* $d(A) = 0$.

If a row is zero, then A doesn't change when we multiply that row by 0. But according to (3.6), $d(A)$ gets multiplied by 0. Thus $d(A) = 0d(A) = 0$.

Rules (3.8′), (3.9′), and (3.6) describe the effect of an elementary row operation (2.7) on the determinant, so they can be rewritten in terms of the elementary matrices. They tell us that $d(EA) = d(A)$ if E is an elementary matrix of the first kind, that $d(EA) = -d(A)$ if E is of the second kind, and (3.6) that $d(EA) = cd(A)$ if E is of the third kind. Let us apply these rules to compute $d(E)$ when E is an elementary matrix. We substitute $A = I$. Then, since $d(I) = 1$, the rules determine $d(EI) = d(E)$:

(3.11) The determinant of an elementary matrix is:

(i) First kind (*add a multiple of one row to another*): $d(E) = 1$, by (3.8′).
(ii) Second kind (*row interchange*): $d(E) = -1$, by (3.9′).
(iii) Third kind (*multiply a row by a nonzero constant*): $d(E) = c$, by (3.6).

Moreover, if we use rules (3.8′), (3.9′), and (3.6) again, applying them this time to an arbitrary matrix A and using the values for $d(E)$ which have just been determined, we obtain the following:

(3.12) *Let E be an elementary matrix and let A be arbitrary. Then*

$$d(EA) = d(E)d(A).$$

Recall from (2.19) that every square matrix A can be reduced by elementary row operations to a matrix A' which is either the identity I or else has its bottom row zero:

$$A' = E_k \cdots E_1 A.$$

We know by (3.5) and (3.10) that $d(A)' = 1$ or $d(A') = 0$ according to the case. By (3.12) and induction,

$$d(A') = d(E_k) \cdots d(E_1) d(A). \tag{3.13}$$

We also know $d(E_i)$, by (3.11), and hence we can use this formula to compute $d(A)$.

(3.14) **Theorem.** *Axiomatic Characterization of the Determinant:* The determinant function (3.4) is the *only* one satisfying rules (3.5–3.7).

Proof. We used only these rules to arrive at equations (3.11) and (3.13), and they determine $d(A)$. Since the expansion by minors (3.4) satisfies (3.5–3.7), it agrees with (3.13). □

We will now return to our usual notation $\det A$ for the determinant of a matrix.

(3.15) **Corollary.** A square matrix A is invertible if and only if $\det A \neq 0$.

This follows from formulas (3.11), (3.13), and (2.18). By (3.11), $\det E_i \neq 0$ for all i. Thus if A' is as in (3.13), then $\det A \neq 0$ if and only if $\det A' \neq 0$, which is the case if and only if $A' = I$. By (2.18), $A' = I$ if and only if A is invertible. □

We can now prove one of the most important properties of the determinant function: its compatibility with matrix multiplications.

(3.16) **Theorem.** Let A, B be any two $n \times n$ matrices. Then

$$\det(AB) = (\det A)(\det B).$$

Proof. We note that this is (3.12) if A is an elementary matrix.

Case 1: A is invertible. By (2.18b), A is a product of elementary matrices: $A = E_1 \cdots E_k$. By (3.12) and induction, $\det A = (\det E_1) \cdots (\det E_k)$, and $\det AB = \det(E_1 \cdots E_k B) = (\det E_1) \cdots (\det E_k)(\det B) = (\det A)(\det B)$.

Case 2: A is not invertible. Then $\det A = 0$ by (3.15), and so the theorem will follow in this case if we show that $\det(AB) = 0$ too. By (2.18), A can be reduced to a matrix $A' = E_k \cdots E_1 A$ having bottom row zero. Then the bottom row of $A'B$ is also zero; hence

$$0 = \det(A'B) = \det(E_k \cdots E_1 AB) = (\det E_k) \cdots (\det E_1)(\det AB).$$

Since $\det E_i \neq 0$, it follows that $\det AB = 0$. □

(3.17) **Corollary.** If A is invertible, $\det(A^{-1}) = \dfrac{1}{\det A}$.

Proof. $(\det A)(\det A^{-1}) = \det I = 1$. □

Note. It is a natural idea to try to define determinants using rules (3.11) and (3.16). These rules certainly determine $\det A$ for every invertible matrix A, since we can write such a matrix as a product of elementary matrices. But there is a problem. Namely, there are many ways to write a given matrix as a product of elementary matrices. Without going through some steps as we have, it is not clear that two such products would give the same answer for the determinant. It is actually not particularly easy to make this idea work.

The proof of the following proposition is a good exercise.

(3.18) **Proposition.** Let A^t denote the transpose of A. Then

$$\det A = \det A^t. \quad \square$$

(3.19) **Corollary.** Properties (3.6–3.10) continue to hold if the word *row* is replaced by *column* throughout. □

4. PERMUTATION MATRICES

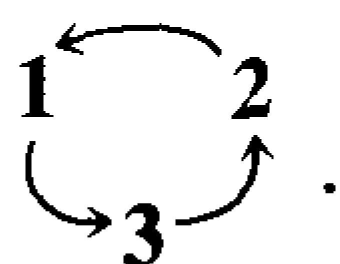
A bijective map p from a set S to itself is called a *permutation* of the set:

$$p\colon S \longrightarrow S. \tag{4.1}$$

For example,

$$\begin{aligned} 1 &\rightsquigarrow 3 \\ 2 &\rightsquigarrow 1 \\ 3 &\rightsquigarrow 2 \end{aligned} \tag{4.2}$$

is a permutation of the set $\{\mathbf{1},\mathbf{2},\mathbf{3}\}$. It is called a *cyclic* permutation because it operates as



There are several notations for permutations. We will use function notation in this section, so that $p(\mathbf{x})$ denotes the value of the permutation p on the element $\mathbf{x}$. Thus if p is the permutation given in (4.2), then

$$p(\mathbf{1}) = \mathbf{3},\ p(\mathbf{2}) = \mathbf{1},\ p(\mathbf{3}) = \mathbf{2}.$$

A *permutation matrix* P is a matrix with the following property: The operation of left multiplication by P is a permutation of the rows of a matrix. The elementary matrices of the second type (2.6ii) are the simplest examples. They correspond to the permutations called *transpositions,* which interchange two rows of a matrix, leaving the others alone. Also,

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \tag{4.3}$$

is a permutation matrix. It acts on a column vector $X = (x, y, z)^t$ as

$$PX = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y \\ z \\ x \end{bmatrix}.$$

The entry in the first position is sent to the third position, and so on, so P has permuted rows according to the cyclic permutation p given in (4.2).

There is one point which can cause confusion and which makes it important for us to establish our notation carefully. When we permute the *entries* of a vector $(x_1,\ldots,x_n)^t$ according to a permutation p, the *indices* are permuted in the opposite way. For instance, multiplying the column vector $X = (x_1, x_2, x_3)^t$ by the matrix in (4.3) gives

$$PX = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_3 \\ x_1 \end{bmatrix}. \tag{4.4}$$

The indices in (4.4) are permuted by $\mathbf{1} \rightsquigarrow \mathbf{2} \rightsquigarrow \mathbf{3} \rightsquigarrow \mathbf{1}$, which is the inverse of the permutation p. Thus there are two ways to associate a permutation to a permutation matrix P: the permutation p which describes how P permutes the entries, and the inverse operation which describes the effect on indices. We must make a decision, so we will say that the permutation associated to P is the one which describes its action on the entries of a column vector. Then the indices are permuted in the opposite way, so

$$PX = \begin{bmatrix} x_{p^{-1}(1)} \\ \vdots \\ x_{p^{-1}(n)} \end{bmatrix}. \tag{4.5}$$

Multiplication by P has the corresponding effect on the rows of an $n \times r$ matrix A.

The permutation matrix P can be written conveniently in terms of the matrix units (2.5) or in terms of certain column vectors called the *standard basis* and denoted by e_i. The vector e_i has a 1 in the ith position as its single nonzero entry, so these vectors are the matrix units for an $n \times 1$ matrix.

(4.6) **Proposition.** Let P be the permutation matrix associated to a permutation p.

(a) The jth column of P is the column vector $e_{p(j)}$.

(b) P is a sum of n matrix units: $P = e_{p(1)1} + \cdots + e_{p(n)n} = \sum_j e_{p(j)j}$. □

A permutation matrix P always has a single 1 in each row and in each column, the rest of its entries being 0. Conversely, any such matrix is a permutation matrix.

(4.7) **Proposition.**

(a) Let p, q be two permutations, with associated permutation matrices P,Q. Then the matrix associated to the permutation pq is the product PQ.

(b) A permutation matrix P is invertible, and its inverse is the transpose matrix: $P^{-1} = P^t$.

Proof. By pq we mean the composition of the two permutations

$$pq(i) = p(q(i)). \tag{4.8}$$

Since P operates by permuting rows according to p and Q operates by permuting according to q, the associative law for matrix multiplication tells us that PQ permutes according to pq:

$$(PQ)X = P(QX).$$

Thus PQ is the permutation matrix associated to pq. This proves (a). We leave the proof of (b) as an exercise. □

The determinant of a permutation matrix is easily seen to be ± 1, using rule (3.9). This determinant is called the *sign of a permutation:*

$$\operatorname{sign} p = \det P = \pm 1. \tag{4.9}$$

The permutation (4.2) has sign $+1$, while any transposition has sign -1 [see (3.11ii)]. A permutation p is called *odd* or *even* according to whether its sign is -1 or $+1$.

Let us now go back to an arbitrary $n \times n$ matrix A and use linearity of the determinant (3.6) to expand $\det A$. We begin by working on the first row. Applying (3.6), we find that

$$\det A = \det\begin{bmatrix} a_{11} 0 \;.\;.\;.\;.\;.\; 0 \\ \text{——} R_2 \text{——} \\ \vdots \\ \text{——} R_n \text{——} \end{bmatrix} + \det\begin{bmatrix} 0 a_{12} 0 \;.\;.\;.\;.\; 0 \\ \text{——} R_2 \text{——} \\ \vdots \\ \text{——} R_n \text{——} \end{bmatrix} + \cdots + \det\begin{bmatrix} 0 \;.\;.\;.\;.\;.\; 0 a_{1n} \\ \text{——} R_2 \text{——} \\ \vdots \\ \text{——} R_n \text{——} \end{bmatrix}.$$

We continue expanding each of these determinants on the second row, and so on. When we are finished, det A is expressed as a sum of many terms, each of which is the determinant of a matrix M having only one entry left in each row:

$$M = \begin{bmatrix} & a_{1?} & \\ a_{2?} & & \\ & & \\ & & a_{n?} \end{bmatrix}.$$

Many of these determinants will be zero because a whole column vanishes. Thus the determinant of a 2×2 matrix is the sum of four terms:

$$\begin{aligned} \det\begin{bmatrix} a & b \\ c & d \end{bmatrix} &= \det\begin{bmatrix} a & 0 \\ c & d \end{bmatrix} + \det\begin{bmatrix} 0 & b \\ c & d \end{bmatrix} \\ &= \det\begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix} + \det\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} + \det\begin{bmatrix} 0 & b \\ c & 0 \end{bmatrix} + \det\begin{bmatrix} 0 & b \\ 0 & d \end{bmatrix}. \end{aligned}$$

But the first and fourth terms are zero; therefore

$$\det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \det\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} + \det\begin{bmatrix} 0 & b \\ c & 0 \end{bmatrix}.$$

In fact, the matrices M having no column zero must have one entry a_{ij} left in each row and each column. They are like permutation matrices P, except that the 1's in P are replaced by the entries of A:

(4.10) $$P = \sum_j e_{p(j)j}, \quad M = \sum_j a_{p(j)j} e_{p(j)j}.$$

By linearity of the determinant (3.6),

$$\begin{aligned} \det M &= (a_{p(1)1} \cdots a_{p(n)n})(\det P) \\ &= (\text{sign } p)(a_{p(1)1} \cdots a_{p(n)n}). \end{aligned}$$

There is one such term for each permutation p. This leads to the formula

(4.11) $$\det A = \sum_{\text{perm}\, p} (\text{sign}\, p) a_{p(1)1} \cdots a_{p(n)1},$$

where the sum is over all permutations of the set $\{\mathbf{1},\ldots,\mathbf{n}\}$. It seems slightly nicer to write this formula in its transposed form:

$$\text{(4.12)} \qquad \det A = \sum_{\text{perm } p} (\text{sign } p) a_{1p(1)} \cdots a_{np(n)}.$$

This is called the *complete expansion* of the determinant.

For example, the complete expansion of the determinant of a 3×3 matrix has six terms:

$$\text{(4.13)} \qquad \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$
$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}.$$

The complete expansion is more of theoretical than of practical importance, because it has too many terms to be useful for computation unless n is small. Its theoretical importance comes from the fact that determinants are exhibited as *polynomials* in the n^2 variable matrix entries a_{ij}, with coefficients ± 1. This has important consequences. Suppose, for example, that each matrix entry a_{ij} is a differentiable function of a single variable: $a_{ij} = a_{ij}(t)$. Then $\det A$ is also a differentiable function of t, because sums and products of differentiable functions are differentiable.

5. CRAMER'S RULE

The name *Cramer's Rule* is applied to a group of formulas giving solutions of systems of linear equations in terms of determinants. To derive these formulas we need to use expansion by minors on columns other than the first one, as well as on rows.

(5.1) *Expansion by minors on the jth column:*

$$\det A = (-1)^{j+1} a_{1j} \det A_{1j} + (-1)^{j+2} a_{2j} \det A_{2j} + \cdots + (-1)^{j+n} a_{nj} \det A_{nj}.$$

(5.2) *Expansion by minors on the ith row:*

$$\det A = (-1)^{i+1} a_{i1} \det A_{i1} + (-1)^{i+2} a_{i2} \det A_{i2} + \cdots + (-1)^{i+n} a_{in} \det A_{in}.$$

In these formulas A_{ij} is the matrix (3.3). The terms $(-1)^{i+j}$ provide alternating signs depending on the position (i, j) in the matrix. (I doubt that such tricky notation is really helpful, but it has become customary.) The signs can be read off of the following figure:

$$\begin{bmatrix} + & - & + & - & \cdots \\ - & + & & & \\ + & - & \cdot & & \\ \vdots & & & \cdot & \\ & & & & \cdot \end{bmatrix}. \tag{5.3}$$

To prove (5.1), one can proceeu in either of two ways:

(a) Verify properties (3.5–3.7) for (5.1) directly and apply Theorem (3.14), or
(b) Interchange (column j) with (column 1) and apply (3.9′) and (3.19).

We omit these verifications. Once (5.1) is proved, (5.2) can be derived from it by transposing the matrix and applying (3.18).

(5.4) **Definition.** Let A be an $n \times n$ matrix. The *adjoint* of A is the $n \times n$ matrix whose (i, j) entry $(\text{adj})_{ij}$ is $(-1)^{i+j} \det A_{ji} = \alpha_{ji}$, where A_{ij} is the matrix obtained by crossing out the ith row and the jth column, as in (3.3):

$$(\text{adj } A) = (\alpha_{ij})^t,$$

where $\alpha_{ij} = (-1)^{i+j} \det A_{ij}$. Thus

$$\text{adj}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{5.5}$$

and

$$\text{adj}\begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 1 & -2 \\ -2 & 0 & 1 \\ -3 & -1 & 2 \end{bmatrix}^t = \begin{bmatrix} 4 & -2 & -3 \\ 1 & 0 & -1 \\ -2 & 1 & 2 \end{bmatrix}. \tag{5.6}$$

We can now proceed to derive the formula called Cramer's Rule.

(5.7) **Theorem.** Let $\delta = \det A$. Then

$$(\text{adj } A) \cdot A = \delta I, \quad \text{and} \quad A \cdot (\text{adj } A) = \delta I.$$

Note that in these equations

$$\delta I = \begin{bmatrix} \delta & & \\ & \ddots & \\ & & \delta \end{bmatrix}.$$

(5.8) **Corollary.** Suppose that the determinant δ of A is not zero. Then

$$A^{-1} = \frac{1}{\delta}(\text{adj } A).$$

For example, the inverse of the 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is

$$\frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

The determinant of the 3×3 matrix whose adjoint is computed in (5.6) happens to be 1; therefore for that matrix, $A^{-1} = \text{adj } A$.

The proof of Theorem (5.7) is easy. The (i, j) entry of $(\text{adj } A) \cdot A$ is

$$(5.9) \qquad (\text{adj})_{i1}a_{1j} + \cdots + (\text{adj})_{in}a_{nj} = \alpha_{1i}a_{1j} + \cdots + \alpha_{ni}a_{nj}.$$

If $i = j$, this is formula (5.1) for δ, which is the required answer. Suppose $i \neq j$. Consider the matrix B obtained by replacing (column i) by (column j) in the matrix A. So (column j) appears twice in the matrix B. Then (5.9) is expansion by minors for B on its ith column. But $\det B = 0$ by (3.7′) and (3.19). So (5.9) is zero, as required. The second equation of Theorem (5.7) is proved similarly. □

Formula (5.8) can be used to write the solution of a system of linear equations $AX = B$, where A is an $n \times n$ matrix in a compact form, provided that $\det A \neq 0$. Multiplying both sides by A^{-1}, we obtain

$$(5.10) \qquad X = A^{-1}B = \frac{1}{\delta}(\text{adj } A)B,$$

where $\delta = \det A$. The product on the right can be expanded out to obtain the formula

$$(5.11) \qquad x_j = \frac{1}{\delta}(b_1\alpha_{1j} + \cdots b_n\alpha_{nj}),$$

where $\alpha_{ij} = \pm\det A_{ij}$ as above.

Notice that the main term $(b_1\alpha_{1j} + \cdots + b_n\alpha_{nj})$ on the right side of (5.11) looks like the expansion of the determinant by minors on the jth column, except that b_i has replaced a_{ij}. We can incorporate this observation to get another expression for the solution of the system of equations. Let us form a new matrix M_j, replacing the jth column of A by the column vector B. Expansion by minors on the jth column shows that

$$\det M_j = (b_1\alpha_{1j} + \cdots + b_n\alpha_{nj}).$$

This gives us the tricky formula

$$(5.12) \qquad x_j = \frac{\det M_j}{\det A}.$$

For some reason it is popular to write the solution of the system of equations $AX = B$ in this form, and it is often this form that is called *Cramer's Rule*. However, this expression does not simplify computation. The main thing to remember is expression (5.8) for the inverse of a matrix in terms of its adjoint; the other formulas follow from this expression.

As with the complete expansion of the determinant (4.10), formulas (5.8–5.11) have theoretical as well as practical significance, because the answers A^{-1} and X are exhibited explicitly as quotients of polynomials in the variables $\{a_{ij}, b_i\}$, with integer coefficients. If, for instance, a_{ij} and b_j are all continuous functions of t, so are the solutions x_i.

A general algebraical determinant in its developed form
may be likened to a mixture of liquids seemingly homogeneous,
but which, being of differing boiling points, admit of being separated
by the process of fractional distillation.

James Joseph Sylvester

EXERCISES

1. The Basic Operations

1. What are the entries a_{21} and a_{23} of the matrix $\begin{bmatrix} 1 & 2 & 5 \\ 2 & 7 & 8 \\ 0 & 9 & 4 \end{bmatrix}$?

2. Compute the products AB and BA for the following values of A and B.

(a) $A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 3 & 1 \end{bmatrix}$, $B = \begin{bmatrix} -8 & -4 \\ 9 & 5 \\ -3 & -2 \end{bmatrix}$

(b) $A = \begin{bmatrix} 1 & 4 \\ 1 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 6 & -4 \\ -3 & 2 \end{bmatrix}$

(c) $A = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}$

3. Let $A = (a_1, \ldots, a_n)$ be a row vector, and let $B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$ be a column vector. Compute the products AB and BA.

4. Verify the associative law for the matrix product

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}.$$

Notice that this is a self-checking problem. You have to multiply correctly, or it won't come out. If you need more practice in matrix multiplication, use this problem as a model.

5. Compute the product $\begin{bmatrix} 1 & a \\ & 1 \end{bmatrix}\begin{bmatrix} 1 & b \\ & 1 \end{bmatrix}$.

6. Compute $\begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}^n$.

7. Find a formula for $\begin{bmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ & & 1 \end{bmatrix}^n$, and prove it by induction.

8. Compute the following matrix products by block multiplication:

$$\left[\begin{array}{cc|cc} 1 & 1 & 1 & 5 \\ 0 & 1 & 0 & 1 \\ \hline 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array}\right]\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ \hline 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 3 \end{array}\right], \quad \left[\begin{array}{c|cc} 0 & 1 & 2 \\ \hline 0 & 1 & 0 \\ 3 & 0 & 1 \end{array}\right]\left[\begin{array}{c|cc} 1 & 2 & 3 \\ \hline 4 & 2 & 3 \\ 5 & 0 & 4 \end{array}\right].$$

9. Prove rule (1.20) for block multiplication.

10. Let A, B be square matrices.
(a) When is $(A + B)(A - B) = A^2 - B^2$?
(b) Expand $(A + B)^3$.

11. Let D be the diagonal matrix

$$\begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix},$$

and let $A = (a_{ij})$ be any $n \times n$ matrix.
(a) Compute the products DA and AD.
(b) Compute the product of two diagonal matrices.
(c) When is a diagonal matrix invertible?

12. An $n \times n$ matrix is called *upper triangular* if $a_{ij} = 0$ whenever $i > j$. Prove that the product of two upper triangular matrices is upper triangular.

13. In each case, find all real 2×2 matrices which commute with the given matrix.
(a) $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$ (e) $\begin{bmatrix} 2 & 3 \\ 0 & 6 \end{bmatrix}$

14. Prove the properties $0 + A = A$, $0A = 0$, and $A0 = 0$ of zero matrices.

15. Prove that a matrix which has a row of zeros is not invertible.

16. A square matrix A is called *nilpotent if* $A^k = 0$ for some $k > 0$. Prove that if A is nilpotent, then $I + A$ is invertible.

17. (a) Find infinitely many matrices B such that $BA = I_2$ when

$$A = \begin{bmatrix} 2 & 3 \\ 1 & 2 \\ 2 & 5 \end{bmatrix}.$$

(b) Prove that there is no matrix C such that $AC = I_3$.

18. Write out the proof of Proposition (1.18) carefully, using the associative law to expand the product $(AB)(B^{-1}A^{-1})$.
19. The *trace* of a square matrix is the sum of its diagonal entries:
$$\operatorname{tr} A = a_{11} + a_{22} + \cdots + a_{nn}.$$
  (a) Show that $\operatorname{tr}(A + B) = \operatorname{tr} A + \operatorname{tr} B$, and that $\operatorname{tr} AB = \operatorname{tr} BA$.
  (b) Show that if B is invertible, then $\operatorname{tr} A = \operatorname{tr} BAB^{-1}$.
20. Show that the equation $AB - BA = I$ has no solutions in $n \times n$ matrices with real entries.

2. Row Reduction

1. (a) For the reduction of the matrix M (2.10) given in the text, determine the elementary matrices corresponding to each operation.
  (b) Compute the product P of these elementary matrices and verify that PM is indeed the end result.
2. Find all solutions of the system of equations $AX = B$ when
$$A = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 3 & 0 & 0 & 4 \\ 1 & -4 & -2 & -2 \end{bmatrix}$$
and B has the following value:
$$\text{(a)} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{(b)} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad \text{(c)} \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}$$
3. Find all solutions of the equation $x_1 + x_2 + 2x_3 - x_4 = 3$.
4. Determine the elementary matrices which are used in the row reduction in Example (2.22) and verify that their product is A^{-1}.
5. Find inverses of the following matrices:
$$\begin{bmatrix} 1 & \\ & 2 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}, \begin{bmatrix} & 1 \\ 1 & \end{bmatrix}, \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}\begin{bmatrix} & 1 \\ 1 & \end{bmatrix}\begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}.$$
6. Make a sketch showing the effect of multiplication by the matrix $A = \begin{bmatrix} 2 & -1 \\ 2 & 3 \end{bmatrix}$ on the plane $\mathbb{R}^2$.
7. How much can a matrix be simplified if both row and column operations are allowed?
8. (a) Compute the matrix product $e_{ij}e_{k\ell}$.
  (b) Write the identity matrix as a sum of matrix units.
  (c) Let A be any $n \times n$ matrix. Compute $e_{ii}Ae_{jj}$.
  (d) Compute $e_{ij}A$ and Ae_{ij}.
9. Prove rules (2.7) for the operations of elementary matrices.
10. Let A be a square matrix. Prove that there is a set of elementary matrices $E_1, \ldots, E_k$ such that $E_k \cdots E_1 A$ either is the identity or has its bottom row zero.
11. Prove that every invertible 2×2 matrix is a product of at most four elementary matrices.
12. Prove that if a product AB of $n \times n$ matrices is invertible then so are the factors A, B.
13. A matrix A is called symmetric if $A = A^t$. Prove that for any matrix A, the matrix AA^t is symmetric and that if A is a square matrix then $A + A^t$ is symmetric.

14. (a) Prove that $(AB)^t = B^tA^t$ and that $A^{tt} = A$.
(b) Prove that if A is invertible then $(A^{-1})^t = (A^t)^{-1}$.

15. Prove that the inverse of an invertible symmetric matrix is also symmetric.

16. Let A and B be symmetric $n \times n$ matrices. Prove that the product AB is symmetric if and only if $AB = BA$.

17. Let A be an $n \times n$ matrix. Prove that the operator "left multiplication by A" determines A in the following sense: If $AX = BX$ for very column vector X, then $A = B$.

18. Consider an arbitrary system of linear equations $AX = B$ where A and B have real entries.
(a) Prove that if the system of equations $AX = B$ has more than one solution then it has infinitely many.
(b) Prove that if there is a solution in the complex numbers then there is also a real solution.

***19.** Prove that the reduced row echelon form obtained by row reduction of a matrix A is uniquely determined by A.

3. Determinants

1. Evaluate the following determinants:

(a) $\begin{bmatrix} 1 & i \\ 2-i & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ (c) $\begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 5 & 2 & 0 & 0 \\ 8 & 6 & 3 & 0 \\ 0 & 9 & 7 & 4 \end{bmatrix}$

(e) $\begin{bmatrix} 1 & 4 & 1 & 3 \\ 2 & 3 & 5 & 0 \\ 4 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{bmatrix}$

2. Prove that $\det \begin{bmatrix} 1 & 2 & 5 & 6 \\ 3 & 1 & 7 & 7 \\ 0 & 0 & 2 & 3 \\ 4 & 2 & 1 & 5 \end{bmatrix} = -\det \begin{bmatrix} 2 & 1 & 5 & 1 \\ 1 & 3 & 7 & 0 \\ 0 & 0 & 2 & 1 \\ 2 & 4 & 1 & 4 \end{bmatrix}$.

3. Verify the rule $\det AB = (\det A)(\det B)$ for the matrices $A = \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 \\ 5 & -2 \end{bmatrix}$. Note that this is a self-checking problem. It can be used as a model for practice in computing determinants.

4. Compute the determinant of the following $n \times n$ matrices by induction on n.

(a) $\begin{bmatrix} & & & 1 \\ & & 1 & \\ & \cdot^{\cdot^{\cdot}} & & \\ 1 & & & \end{bmatrix}$ (b) $\begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & -1 & \ddots & & \\ & & & & 2 & -1 \\ & & & & -1 & 2 \end{bmatrix}$

5. Evaluate $\det \begin{bmatrix} 1 & 2 & 3 & \cdots & n \\ 2 & 2 & 3 & & \vdots \\ 3 & 3 & 3 & & \vdots \\ \vdots & & & \ddots & \vdots \\ n & & \cdots & & n \end{bmatrix}$

*6. Compute $\det \begin{bmatrix} 2 & 1 & & & & & & \\ 1 & 2 & 1 & & & & & \\ & 1 & 2 & 1 & & & & \\ & & 1 & 2 & 1 & & & \\ & & & 1 & 2 & 1 & & 1 \\ & & & & 1 & 2 & 1 & \\ & & & & & 1 & 2 & \\ & & & & 1 & & & 2 \end{bmatrix}$.

7. Prove that the determinant is linear in the rows of a matrix, as asserted in (3.6).
8. Let A be an $n \times n$ matrix. What is $\det(-A)$?
9. Prove that $\det A^t = \det A$.
10. Derive the formula $\det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$ from the properties (3.5, 3.6, 3.7, 3.9).
11. Let A and B be square matrices. Prove that $\det(AB) = \det(BA)$.
12. Prove that $\det\begin{bmatrix} A & B \\ 0 & D \end{bmatrix} = (\det A)(\det D)$, if A and D are square blocks.

*13. Let a $2n \times 2n$ matrix be given in the form $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, where each block is an $n \times n$ matrix. Suppose that A is invertible and that $AC = CA$. Prove that $\det M = \det(AD - CB)$. Give an example to show that this formula need not hold when $AC \neq CA$.

4. Permutation Matrices

1. Consider the permutation p defined by $\mathbf{1} \rightsquigarrow \mathbf{3}$, $\mathbf{2} \rightsquigarrow \mathbf{1}$, $\mathbf{3} \rightsquigarrow \mathbf{4}$, $\mathbf{4} \rightsquigarrow \mathbf{2}$.
   (a) Find the associated permutation matrix P.
   (b) Write p as a product of transpositions and evaluate the corresponding matrix product.
   (c) Compute the sign of p.
2. Prove that every permutation matrix is a product of transpositions.
3. Prove that every matrix with a single 1 in each row and a single 1 in each column, the other entries being zero, is a permutation matrix.
4. Let p be a permutation. Prove that $\operatorname{sign} p = \operatorname{sign} p^{-1}$.
5. Prove that the transpose of a permutation matrix P is its inverse.
6. What is the permutation matrix associated to the permutation $\mathbf{i} \rightsquigarrow \mathbf{n-i}$?
7. (a) The complete expansion for the determinant of a 3×3 matrix consists of six triple products of matrix entries, with sign. Learn which they are.
   (b) Compute the determinant of the following matrices using the complete expansion, and check your work by another method:

$$\begin{bmatrix} 1 & 1 & 2 \\ 2 & 4 & 2 \\ 0 & 2 & 1 \end{bmatrix}, \begin{bmatrix} 4 & -1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 1 \end{bmatrix}, \begin{bmatrix} a & b & c \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

8. Prove that the complete expansion (4.12) defines the determinant by verifying rules (3.5–3.7).
9. Prove that formulas (4.11) and (4.12) define the same number.

5. Cramer's Rule

1. Let $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a matrix with determinant 1. What is A^{-1}?

2. (self-checking) Compute the adjoints of the matrices $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $\begin{bmatrix} 1 & 1 & 2 \\ 2 & 4 & 2 \\ 0 & 2 & 1 \end{bmatrix}$, $\begin{bmatrix} 4 & -1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 1 \end{bmatrix}$, and $\begin{bmatrix} a & b & c \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$, and verify Theorem (5.7) for them.

3. Let A be an $n \times n$ matrix with integer entries a_{ij}. Prove that A^{-1} has integer entries if and only if $\det A = \pm 1$.

4. Prove that expansion by minors on a row of a matrix defines the determinant function.

Miscellaneous Problems

1. Write the matrix $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ as a product of elementary matrices, using as few as you can. Prove that your expression is as short as possible.

2. Find a representation of the complex numbers by real 2×2 matrices which is compatible with addition and multiplication. Begin by finding a nice solution to the matrix equation $A^2 = -I$.

3. (*Vandermonde determinant*) **(a)** Prove that $\det \begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{bmatrix} = (b - a)(c - a)(c - b)$.

*__(b)__ Prove an analogous formula for $n \times n$ matrices by using row operations to clear out the first column cleverly.

***4.** Consider a general system $AX = B$ of m linear equations in n unknowns. If the coefficient matrix A has a left inverse A', a matrix such that $A'A = I_n$, then we may try to solve the system as follows:

$$AX = B$$

$$A'AX = A'B$$

$$X = A'B.$$

But when we try to check our work by running the solution backward, we get into trouble:

$$X = A'B$$

$$AX = AA'B$$

$$AX \stackrel{?}{=} B.$$

We seem to want A' to be a right inverse: $AA' = I_m$, which isn't what was given. Explain. (Hint: Work out some examples.)

5. **(a)** Let A be a real 2×2 matrix, and let A_1, A_2 be the rows of A. Let P be the parallelogram whose vertices are $0, A_1, A_2, A_1 + A_2$. Prove that the area of P is the absolute value of the determinant $\det A$ by comparing the effect of an elementary row operation on the area and on $\det A$.

***(b)** Prove an analogous result for $n \times n$ matrices.

***6.** Most invertible matrices can be written as a product $A = LU$ of a lower triangular matrix L and an upper triangular matrix U, where in addition all diagonal entries of U are 1.

(a) Prove uniqueness, that is, prove that there is at most one way to write A as a product.

(b) Explain how to compute L and U when the matrix A is given.

(c) Show that every invertible matrix can be written as a product LPU, where L, U are as above and P is a permutation matrix.

7. Consider a system of n linear equations in n unknowns: $AX = B$, where A and B have *integer* entries. Prove or disprove the following.

(a) The system has a rational solution if $\det A \neq 0$.

(b) If the system has a rational solution, then it also has an integer solution.

***8.** Let A, B be $m \times n$ and $n \times m$ matrices. Prove that $I_m - AB$ is invertible if and only if $I_n - BA$ is invertible.

Chapter 2

Groups

Il est peu de notions en mathématiques qui soient plus primitives que celle de loi de composition.

Nicolas Bourbaki

1. THE DEFINITION OF A GROUP

In this chapter we study one of the most important algebraic concepts, that of a *group*. A group is a set on which a law of composition is defined, such that all elements have inverses. The precise definition is given below in (1.10). For example, the set of nonzero real numbers forms a group $\mathbb{R}^\times$ under multiplication, and the set of all real numbers forms a group $\mathbb{R}^+$ under addition. The set of invertible $n \times n$ matrices, called the general linear group, is a very important example in which the law of composition is matrix multiplication. We will see many more examples as we go along.

By a *law of composition* on a set S, we mean a rule for combining pairs a, b of elements S to get another element, say p, of S. The original models for this notion are addition and multiplication of real numbers. Formally, a law of composition is a function of two variables on S, with values in S, or it is a map

$$S \times S \longrightarrow S$$

$$a, b \rightsquigarrow p.$$

Here, $S \times S$ denotes, as always, the product set of pairs (a, b) of elements of S.

Functional notation $p = f(a, b)$ isn't very convenient for laws of composition. Instead, the element obtained by applying the law to a pair (a, b) is usually denoted using a notation resembling those used for multiplication or addition:

$$p = ab,\ a \times b,\ a \circ b,\ a + b, \text{ and so on,}$$

a choice being made for the particular law in question. We call the element p the *product* or *sum* of a and b, depending on the notation chosen.

Our first example of a law of composition, and one of the two main examples, is matrix multiplication on the set S of $n \times n$ matrices.

We will use the product notation ab most frequently. Anything we prove with product notation can be rewritten using another notation, such as addition. It will continue to be valid, because the rewriting is just a change of notation.

It is important to note that the symbol ab is a notation for a certain element of S. Namely, it is the element obtained by applying the given law of composition to the elements called a and b. Thus if the law is multiplication of matrices and if $a = \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}$ and $b = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$, then ab denotes the matrix $\begin{bmatrix} 7 & 3 \\ 4 & 2 \end{bmatrix}$. Once the product ab has been evaluated, the elements a and b can not be recovered from it.

Let us consider a law of composition written multiplicatively as ab. It will be called *associative* if the rule

$$(1.1) \qquad (ab)c = a(bc) \quad \textit{(associative law)}$$

holds for all a, b, c in S, and *commutative* if

$$(1.2) \qquad ab = ba \quad \textit{(commutative law)}$$

holds for all a, b in S. Our example of matrix multiplication is associative but not commutative.

When discussing groups in general, we will use multiplicative notation. It is customary to reserve additive notation $a + b$ for commutative laws of composition, that is, when $a + b = b + a$ for all a, b. Multiplicative notation carries no implication either way concerning commutativity.

In additive notation the associative law is $(a + b) + c = a + (b + c)$, and in functional notation it is

$$f(f(a,b),c) = f(a, f(b,c)).$$

This ugly formula illustrates the fact that functional notation isn't convenient for algebraic manipulation.

The associative law is more fundamental than the commutative law; one reason for this is that composition of functions, our second example of a law of composition, is associative. Let T be a set, and let g, f be functions (or maps) from T to T. Let $g \circ f$ denote the composed map $t \rightsquigarrow g(f(t))$. The rule

$$g, f \rightsquigarrow g \circ f$$

is a law of composition on the set $S = \text{Maps}(T, T)$ of all maps $T \longrightarrow T$.

As is true for matrix multiplication, composition of functions is an associative law. For if f, g, h are three maps from T to itself, then $(h \circ g) \circ f = h \circ (g \circ f)$:

$$T \xrightarrow{f} T \xrightarrow{g} T \xrightarrow{h} T.$$

(with arrows $h \circ g$ from the second T to the fourth T, and $g \circ f$ from the first T to the third T)

This is clear, since both of the composed maps send $t \rightsquigarrow h(g(f(t)))$.

The simplest example is that T is a set of two elements $\{a, b\}$. Then there are four maps $T \longrightarrow T$:

i: the *identity* map, defined by $i(a) = a$, $i(b) = b$;
τ: the *transposition*, defined by $\tau(a) = b$, $\tau(b) = a$;
α: the constant function $\alpha(a) = \alpha(b) = a$;
β: the constant function $\beta(a) = \beta(b) = b$.

The law of composition on S can be exhibited in a *multiplication table* as follows:

(1.3)

	i	τ	α	β
i	i	τ	α	β
τ	τ	i	β	α
α	α	α	α	α
β	β	β	β	β

which is to be read in this way:

	$\cdots\; v \;\cdots$
$\vdots$	$\vdots$
u	$\cdots\; u \circ v$
$\vdots$	

Thus $\tau \circ \alpha = \beta$, while $\alpha \circ \tau = \alpha$. Composition of functions is not commutative.

Going back to a general law of composition, suppose we want to define the product of a string of n elements of a set:

$$a_1 a_2 \cdots a_n = ?$$

There are various ways to do this using the given law, which tells us how to multiply two elements. For instance, we could first use the law to find the product $a_1 a_2$, then multiply this element by a_3, and so on:

$$((a_1 a_2) a_3) a_4 \cdots.$$

When $n = 4$, there are four other ways to combine the same elements; $(a_1 a_2)(a_3 a_4)$ is one of them. It can be proved by induction that if the law is *associative,* then all such products are equal. This allows us to speak of the product of an arbitrary string of elements.

(1.4) **Proposition.** Suppose an associative law of composition is given on a set S. There is a unique way to define, for every integer n, a product of n elements $a_1, \ldots, a_n$ of S (we denote it temporarily by $[a_1 \cdots a_n]$) with the following properties:

(i) the product $[a_1]$ of one element is the element itself;
(ii) the product $[a_1a_2]$ of two elements is given by the law of composition;
(iii) for any integer i between 1 and n, $[a_1 \cdots a_n] = [a_1 \cdots a_i][a_{i+1} \cdots a_n]$.

The right side of equation (iii) means that the two products $[a_1 \cdots a_i]$ and $[a_{i+1} \cdots a_n]$ are formed first and the results are then multiplied using the given law of composition.

Proof. We use induction on n. The product is defined by (i) and (ii) for $n \leq 2$, and it does satisfy (iii) when $n = 2$. Suppose that we know how to define the product of r elements when $r \leq n - 1$, and that this product is the unique product satisfying (iii). We then define the product of n elements by the rule

$$[a_1 \cdots a_n] = [a_1 \cdots a_{n-1}][a_n],$$

where the terms on the right side are those already defined. If a product satisfying (iii) exists, then this formula gives the product because it is (iii) when $i = n - 1$. So if it exists, the product is unique. We must now check (iii) for $i < n - 1$:

$$\begin{aligned}
[a_1 \cdots a_n] &= [a_1 \cdots a_{n-1}][a_n] && \textit{(our definition)} \\
&= ([a_1 \cdots a_i][a_{i+1} \cdots a_{n-1}])[a_n] && \textit{(induction hypothesis)} \\
&= [a_1 \cdots a_i]([a_{i+1} \cdots a_{n-1}][a_n]) && \textit{(associative law)} \\
&= [a_1 \cdots a_i][a_{i+1} \cdots a_n] && \textit{(induction hypothesis)}.
\end{aligned}$$

This completes the proof. We will drop the brackets from now on and denote the product by $a_1 \cdots a_n$. □

An *identity* for a law of composition is an element e of S having the property that

(1.5) $$ea = a \quad \text{and} \quad ae = a, \text{ for all } a \in S.$$

There can be at most one identity element. For if e, e' were two such elements, then since e is an identity, $ee' = e'$, and since e' is an identity, $ee' = e$. Thus $e = e'$.

Both of our examples, matrix multiplication and composition of functions, have an identity. For $n \times n$ matrices it is the identity matrix I, and for $\text{Maps}(T, T)$ it is the identity map, which carries each element of T to itself.

Often the identity is denoted by 1 if the law of composition is written multiplicatively, or by 0 if it is written additively. These elements do not need to be related to the *numbers* 1 and 0, but they share the property of being identity elements for their laws of composition.

Suppose that our law of composition has an identity, and let us use the symbol 1 for it. An element $a \in S$ is called *invertible* if there is another element b such that

$$ab = 1 \quad \text{and} \quad ba = 1.$$

As with matrix multiplication [Chapter 1 (1.17)], it follows from the associative law that the inverse is unique if it exists. It is denoted by a^{-1}:

$$aa^{-1} = a^{-1}a = 1.$$

Inverses multiply in the opposite order:

$$(ab)^{-1} = b^{-1}a^{-1}. \tag{1.6}$$

The proof is the same as for matrices [Chapter 1 (1.18)].

Power notation may be used for an associative law of composition:

$$\begin{aligned} a^n &= \underbrace{a \cdots a}_{n \text{ times}} && (n \geq 1) \\ a^0 &= 1 && \text{provided the identity exists} \\ a^{-n} &= a^{-1} \cdots a^{-1} && \text{provided } a \text{ is invertible.} \end{aligned} \tag{1.7}$$

The usual rules for manipulation of powers hold:

$$a^{r+s} = a^r a^s \quad \text{and} \quad (a^r)^s = a^{rs}. \tag{1.8}$$

It isn't advisable to introduce fraction notation

$$\frac{b}{a} \tag{1.9}$$

unless the law of composition is commutative, for it is not clear from the notation whether the fraction stands for ba^{-1} or $a^{-1}b$, and these two elements may be different.

When additive notation is used for the law of composition, the inverse is denoted by $-a$, and the power notation a^n is replaced by the notation $na = a + \cdots + a$, as with addition of real numbers.

(1.10) **Definition.** A *group* is a set G together with a law of composition which is associative and has an identity element, and such that every element of G has an inverse.

It is customary to denote the group and the set of its elements by the same symbol.

An *abelian group* is a group whose law of composition is commutative. Additive notation is often used for abelian groups. Here are some simple examples of abelian groups:

(1.11)
- $\mathbb{Z}^+$: the integers, with addition;
- $\mathbb{R}^+$: the real numbers, with addition;
- $\mathbb{R}^\times$: the nonzero real numbers, with multiplication;
- $\mathbb{C}^+, \mathbb{C}^\times$: the analogous groups, where the set $\mathbb{C}$ of complex numbers replaces the real numbers $\mathbb{R}$.

Here is an important property of groups:

(1.12) **Proposition.** *Cancellation Law:* Let a, b, c be elements of a group G. If $ab = ac$, then $b = c$. If $ba = ca$, then $b = c$.

Proof. Multiply both sides of $ab = ac$ by a^{-1} on the left: $b = a^{-1}ab = a^{-1}ac = c$. □

Multiplication by a^{-1} in this proof is not a trick; it is essential. If an element a is not invertible, the cancellation law need not hold. For instance, $0 \cdot 1 = 0 \cdot 2$, or

$$\begin{bmatrix} 1 & \\ & \end{bmatrix}\begin{bmatrix} 1 & \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & \\ & \end{bmatrix}\begin{bmatrix} 1 & \\ & 3 \end{bmatrix}.$$

The two most basic examples of groups are obtained from the examples of laws of composition that we have considered—multiplication of matrices and composition of functions—by leaving out the elements which are not invertible. As we remarked in Chapter 1, the $n \times n$ *general linear group* is the group of all invertible $n \times n$ matrices. It is denoted by

$$GL_n = \{n \times n \text{ matrices } A \text{ with } \det A \neq 0\}. \tag{1.13}$$

If we want to indicate that we are working with real or complex matrices, we write

$$GL_n(\mathbb{R}) \quad \text{or} \quad GL_n(\mathbb{C}),$$

according to the case.

In the set $S = \text{Maps}(T, T)$ of functions, a map $f: T \longrightarrow T$ has an inverse function if and only if it is bijective. Such a map is also called a *permutation* of T. The set of permutations forms a group. In Example (1.3), the invertible elements are i and τ, and they form a group with two elements. These two elements are the permutations of the set $\{a, b\}$.

The group of permutations of the set $\{\mathbf{1},\mathbf{2},\ldots,\mathbf{n}\}$ of integers from 1 to n is called the *symmetric group* and is denoted by S_n:

$$S_n = \text{group of permutations of } \{\mathbf{1},\ldots,\mathbf{n}\}. \tag{1.14}$$

Because there are $n!$ permutations of a set of n elements, this group contains $n!$ elements. (We say that the *order* of the group is $n!$.) The symmetric group S_2 consists of the two elements i and τ, where i denotes the identity permutation and τ denotes the transposition which interchanges $\mathbf{1},\mathbf{2}$ as in (1.3). The group law, composition of functions, is described by the fact that i is the identity element and by the relation $\tau\tau = \tau^2 = i$.

The structure of S_n becomes complicated very rapidly as n increases, but we can work out the case $n = 3$ fairly easily. The symmetric group S_3 contains six elements. It will be an important example for us because it is the smallest group whose law of composition is not commutative. To describe this group, we pick two particular permutations x,y in terms of which we can write all others. Let us take for x the cyclic permutation of the indices. It is represented by matrix (4.3) from Chapter 1:

$$x = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}. \tag{1.15}$$

For y, we take the transposition which interchanges **1**, **2**, fixing **3**:

(1.16)
$$y = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The six permutations of $\{\mathbf{1},\mathbf{2},\mathbf{3}\}$ are

(1.17)
$$\{1, x, x^2, y, xy, x^2y\} = \{x^iy^j \mid 0 \le i \le 2, 0 \le j \le 1\},$$

where 1 denotes the identity permutation. This can be verified by computing the products.

The rules

(1.18)
$$x^3 = 1,\ y^2 = 1,\ yx = x^2y$$

can also be verified directly. They suffice for computation in the group S_3. Any product of the elements x,y and of their inverses, such as $x^{-1}y^3x^2y$ for instance, can be brought into the form x^iy^j with $0 \le i \le 2$ and $0 \le j \le 1$ by applying the above rules repeatedly. To do so, we move all occurrences of y to the right side using the last relation and bring the exponents into the indicated ranges using the first two relations:

$$x^{-1}y^3x^2y = x^2yx^2y = x^2(yx)xy = x^2x^2yxy = \cdots = x^6y^2 = 1.$$

Therefore one can write out a complete multiplication table for S_3 with the aid of these rules. Because of this, the rules are called *defining relations* for the group, a concept which we will study formally in Chapter 6.

Note that the commutative law does not hold in S_3, because $yx \neq xy$.

2. SUBGROUPS

One reason that the general linear group and the symmetric group are so important is that many other groups are contained in them as subgroups. A subset H of a group G is called a *subgroup* if it has the following properties:

(2.1) (a) *Closure:* If $a \in H$ and $b \in H$, then $ab \in H$.
(b) *Identity:* $1 \in H$.
(c) *Inverses:* If $a \in H$, then $a^{-1} \in H$.

These conditions are explained as follows: The first condition (a) tells us that the law of composition on the group G can be used to define a law on H, called the *induced law of composition*. The second and third conditions (b, c) say that H is a group with respect to this induced law. Notice that (2.1) mentions all parts of the definition of a group except for the associative law. We do not need to mention associativity. It carries over automatically from G to H.

Every group has two obvious subgroups: the whole group and the subgroup $\{1\}$ consisting of the identity element alone. A subgroup is said to be a *proper subgroup* if it is not one of these two.

Here are two examples of subgroups:

(2.2) **Examples.**

(a) The set T of invertible upper triangular 2×2 matrices

$$\begin{bmatrix} a & b \\ & d \end{bmatrix} \qquad (a, d \neq 0)$$

is a subgroup of the general linear group $GL_2(\mathbb{R})$.

(b) The set of complex numbers of absolute value 1—the set of points on the unit circle in the complex plane—is a subgroup of $\mathbb{C}^\times$.

As a further example, we will determine the subgroups of the additive group $\mathbb{Z}^+$ of integers. Let us denote the subset of $\mathbb{Z}$ consisting of all multiples of a given integer b by $b\mathbb{Z}$:

$$b\mathbb{Z} = \{n \in \mathbb{Z} \mid n = bk \text{ for some } k \in \mathbb{Z}\}. \tag{2.3}$$

(2.4) **Proposition.** For any integer b, the subset $b\mathbb{Z}$ is a subgroup of $\mathbb{Z}^+$. Moreover, every subgroup H of $\mathbb{Z}^+$ is of the type $H = b\mathbb{Z}$ for some integer b.

Proof. We leave the verification that $b\mathbb{Z}$ is a subgroup as an exercise and proceed to show that every subgroup has this form. Let H be a subgroup of $\mathbb{Z}^+$. Remember that the law of composition on $\mathbb{Z}^+$ is addition, the identity element is 0, and the inverse of a is $-a$. So the axioms for a subgroup read

(i) if $a \in H$ and $b \in H$, then $a + b \in H$;

(ii) $0 \in H$;

(iii) if $a \in H$, then $-a \in H$.

By axiom (ii), $0 \in H$. If 0 is the only element of H, then $H = 0\mathbb{Z}$, so that case is settled. If not, there is a positive integer in H. For let $a \in H$ be any nonzero element. If a is negative, then $-a$ is positive, and axiom (iii) tells us that $-a$ is in H. We choose for b the smallest positive integer in H, and we claim that $H = b\mathbb{Z}$. We first show that $b\mathbb{Z} \subset H$, in other words, that $bk \in H$ for every integer k. If k is a positive integer, then $bk = b + b + \cdots + b$ (k terms). This element is in H by axiom (i) and induction. So is $b(-k) = -bk$, by axiom (iii). Finally, axiom (ii) tells us that $b0 = 0 \in H$.

Next we show that $H \subset b\mathbb{Z}$, that is, that every element $n \in H$ is an integer multiple of b. We use division with remainder to write $n = bq + r$, where q, r are integers and where the remainder r is in the range $0 \leq r < b$. Then n and bq are both in H, and axioms (iii) and (i) show that $r = n - bq$ is in H too. Now by our

choice, b is the smallest positive integer in H, while $0 \le r < b$. Therefore $r = 0$, and $n = bq \in b\mathbb{Z}$, as required. □

The elements of the subgroup $b\mathbb{Z}$ can be described as the integers which are divisible by b. This description leads to a striking application of proposition (2.3) to subgroups which are generated by *two* integers a, b. Let us assume that a and b are not both zero. The set

$$(2.5) \qquad a\mathbb{Z} + b\mathbb{Z} = \{n \in \mathbb{Z} \mid n = ar + bs \text{ for some integers } r, s\}$$

is a subgroup of $\mathbb{Z}^+$. It is called the subgroup *generated* by a and b, because it is the smallest subgroup which contains both of these elements. Proposition (2.3) tells us that this subgroup has the form $d\mathbb{Z}$ for some integer d, so it is the set of integers which are divisible by d. The generator d is called the *greatest common divisor* of a and b, for reasons which are explained in the following proposition:

(2.6) **Proposition.** Let a, b be integers, not both zero, and let d be the positive integer which generates the subgroup $a\mathbb{Z} + b\mathbb{Z}$. Then

(a) d can be written in the form $d = ar + bs$ for some integers r and s.
(b) d divides a and b.
(c) If an integer e divides a and b, it also divides d.

Proof. The first assertion (a) just restates the fact that d is contained in $a\mathbb{Z} + b\mathbb{Z}$. Next, notice that a and b are in the subgroup $d\mathbb{Z} = a\mathbb{Z} + b\mathbb{Z}$. Therefore d divides a and b. Finally, if e is an integer which divides a and b, then a and b are in $e\mathbb{Z}$. This being so, any integer $n = ar + bs$ is also in $e\mathbb{Z}$. By assumption, d has this form, so e divides d. □

If two integers a, b are given, one way to find their greatest common divisor is to factor each of them into prime integers and then collect the common ones. Thus the greatest common divisor of $36 = 2\cdot 2\cdot 3\cdot 3$ and $60 = 2\cdot 2\cdot 3\cdot 5$ is $12 = 2\cdot 2\cdot 3$. Properties (2.6ii, iii) are easy to verify. But without proposition (2.4), the fact that the integer determined by this method has the form $ar + bs$ would not be clear at all. (In our example, $12 = 36\cdot 2 - 60\cdot 1$.) We will discuss the applications of this fact to arithmetic in Chapter 11.

We now come to an important abstract example of a subgroup, the *cyclic subgroup* generated by an arbitrary element x of a group G. We use multiplicative notation. The cyclic subgroup H generated by x is the set of all powers of x:

$$(2.7) \qquad H = \{\ldots, x^{-2}, x^{-1}, 1, x, x^2, \ldots\}.$$

It is a subgroup of G—the smallest subgroup which contains x. But to interpret (2.7) correctly, we must remember that x^n is a notation for a certain element of G. It may happen that there are repetitions in the list. For example, if $x = 1$, then all elements in the list are equal to 1. We may distinguish two possibilities: Either the powers of

x are all distinct elements, or they are not. In the first case, the group H is called *infinite cyclic*.

Suppose we have the second case, so that two powers are equal, say $x^n = x^m$, where $n > m$. Then $x^{n-m} = 1$ [Cancellation Law (1.12)], and so there is a nonzero power of x which is equal to 1.

(2.8) **Lemma.** The set S of integers n such that $x^n = 1$ is a subgroup of $\mathbb{Z}^+$.

Proof. If $x^m = 1$ and $x^n = 1$, then $x^{m+n} = x^m x^n = 1$ too. This shows that $m + n \in S$ if $m, n \in S$. So axiom (i) for a subgroup is verified. Also, axiom (ii) holds because $x^0 = 1$. Finally, if $x^n = 1$, then $x^{-n} = x^n x^{-n} = x^0 = 1$. Thus $-n \in S$ if $n \in S$. □

It follows from Lemma (2.8) and Proposition (2.4) that $S = m\mathbb{Z}$, where m is the smallest positive integer such that $x^m = 1$. The m elements $1, x, \ldots, x^{m-1}$ are all different. (If $x^i = x^j$ with $0 \le i < j < m$, then $x^{j-i} = 1$. But $j - i < m$, so this is impossible.) Moreover, any power x^n is equal to one of them: By division with remainder, we may write $n = mq + r$ with remainder r less than m. Then $x^n = (x^m)^q x^r = x^r$. Thus H consists of the following m elements:

(2.9) $$H = \{1, x, \ldots, x^{m-1}\}, \textit{ these powers are distinct, and } x^m = 1.$$

Such a group is called a *cyclic group of order m*.

The *order* of any group G is the number of its elements. We will often denote the order by

(2.10) $$|G| = \textit{number of elements of } G.$$

Of course, the order may be infinite.

An element of a group is said to have *order m* (possibly infinity) if the cyclic subgroup it generates has order m. This means that m is the smallest positive integer with the property $x^m = 1$ or, if the order is infinite, that $x^m \neq 1$ for all $m \neq 0$.

For example, the matrix $\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$ is an element of order 6 in $GL_2(\mathbb{R})$, so the cyclic subgroup it generates has order 6. On the other hand, the matrix $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ has infinite order, because

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}.$$

We may also speak of the subgroup of a group G *generated by a subset U*. This is the smallest subgroup of G containing U, and it consists of all elements of G which can be expressed as a product of a string of elements of U and of their inverses. In particular, a subset U of G is said to *generate* G if every element of G is such a product. For example, we saw in (1.17) that the set $U = \{x, y\}$ generates the symmetric group S_3. Proposition (2.18) of Chapter 1 shows that the elementary matrices generate GL_n.

The *Klein four group* V is the simplest group which is not cyclic. It will appear in many forms. For instance, it can be realized as the group consisting of the four matrices

$$(2.11)\qquad \begin{bmatrix} \pm 1 & \\ & \pm 1 \end{bmatrix}.$$

Any two elements different from the identity generate V.

The *quaternion group* H is another example of a small subgroup of $GL_2(\mathbb{C})$ which is not cyclic. It consists of the eight matrices

$$(2.12)\qquad H = \{\pm\mathbf{1}, \pm\mathbf{i}, \pm\mathbf{j}, \pm\mathbf{k}\},$$

where

$$\mathbf{1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{i} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \mathbf{j} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \mathbf{k} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

The two elements $\mathbf{i}, \mathbf{j}$ generate H, and computation leads to the formulas

$$(2.13)\qquad \mathbf{i}^4 = 1, \quad \mathbf{i}^2 = \mathbf{j}^2, \mathbf{ji} = \mathbf{i}^3\mathbf{j}.$$

These products determine the multiplication table of H.

3. ISOMORPHISMS

Let G and G' be two groups. We want to say that they are *isomorphic* if all properties of the group structure of G hold for G' as well, and conversely. For example, let G be the set of real matrices of the form

$$\begin{bmatrix} 1 & x \\ & 1 \end{bmatrix}.$$

This is a subgroup of $GL_2(\mathbb{R})$, and the product of two such matrices is

$$\begin{bmatrix} 1 & x \\ & 1 \end{bmatrix}\begin{bmatrix} 1 & y \\ & 1 \end{bmatrix} = \begin{bmatrix} 1 & x+y \\ & 1 \end{bmatrix}.$$

The upper right entries of the matrices add when the matrices are multiplied, the rest of the matrix being fixed. So when computing with such matrices, we need to keep track of only the upper right entry. This fact is expressed formally by saying that the group G is isomorphic to the additive group of real numbers.

How to make the concept of isomorphism precise will not be immediately clear, but it turns out that the right way is to relate two groups by a *bijective correspondence* between their elements, *compatible with the laws of composition*, that is, a correspondence

$$(3.1)\qquad G \longleftrightarrow G'$$

having this property: If $a, b \in G$ correspond to $a', b' \in G'$, then the product ab in G corresponds to the product $a'b'$ in G'. When this happens, all properties of the group structure carry over from one group to the other.

For example, the identity elements in isomorphic groups G and G' correspond. To see this, say that the identity element 1 of G corresponds to an element ϵ' in G'. Let a' be an arbitrary element of G', and let a be the corresponding element of G. By assumption, products correspond to products. Since $1a = a$ in G, it follows that $\epsilon' a' = a'$ in G'. In this way, one shows that $\epsilon' = 1'$. Another example: The orders of corresponding elements are equal. If a corresponds to a' in G', then, since the correspondence is compatible with multiplication, $a^r = 1$ if and only if $a'^r = 1'$.

Since two isomorphic groups have the same properties, it is often convenient to identify them with each other when speaking informally. For example, the symmetric group S_n of permutations of $\{\mathbf{1},\ldots,\mathbf{n}\}$ is isomorphic to the group of permutation matrices, a subgroup of $GL_n(\mathbb{R})$, and we often blur the distinction between these two groups.

We usually write the correspondence (3.1) asymmetrically as a function, or map $\varphi: G \longrightarrow G'$. Thus an *isomorphism* φ *from* G *to* G' is a bijective map which is compatible with the laws of composition. If we write out what this compatibility means using function notation for φ, we get the condition

(3.2) $$\varphi(ab) = \varphi(a)\varphi(b), \text{ for all } a, b \in G.$$

The left side of this equality means to multiply a and b in G and then apply φ, while on the right the elements $\varphi(a)$ and $\varphi(b)$, which we denoted by a', b' before, are multiplied in G'. We could also write this condition as

$$(ab)' = a'b'.$$

Of course, the choice of G as domain for this isomorphism is arbitrary. The inverse function $\varphi^{-1}: G' \longrightarrow G$ would serve just as well.

Two groups G and G' are called *isomorphic* if there exists an isomorphism $\varphi: G \longrightarrow G'$. We will sometimes indicate that two groups are isomorphic by the symbol $\approx$:

(3.3) $$G \approx G' \text{ means } G \text{ is isomorphic to } G'.$$

For example, let $C = \{\ldots, a^{-2}, a^{-1}, 1, a, a^2, \ldots\}$ be an infinite cyclic group. Then the map

$$\varphi: \mathbb{Z}^+ \longrightarrow C$$

defined by $\varphi(n) = a^n$ is an isomorphism. Since the notation is additive in the domain and multiplicative in the range, condition (3.2) translates in this case to $\varphi(m + n) = \varphi(m)\varphi(n)$, or

$$a^{m+n} = a^m a^n.$$

One more simple example:
Let $G = \{1, x, x^2, \ldots, x^{n-1}\}$ and $G' = \{1, y, y^2, \ldots, y^{n-1}\}$ be two cyclic groups, generated by elements x, y of the same order. Then the map which sends x^i to y^i is an isomorphism: Two cyclic groups of the same order are isomorphic.

Recapitulating, two groups G and G' are isomorphic if there exists an isomorphism $\varphi: G \longrightarrow G'$, a bijective map compatible with the laws of composition. The groups isomorphic to a given group G form what is called the *isomorphism class* of G, and any two groups in an isomorphism class are isomorphic. When one speaks of *classifying* groups, what is meant is to describe the isomorphism classes. This is too hard to do for all groups, but we will see later that there is, for example, one isomorphism class of groups of order 3 [see (6.13)], and that there are two classes of groups of order 4 and five classes of groups of order 12 [Chapter 6 (5.1)].

A confusing point about isomorphisms is that there exist isomorphisms from a group G to itself:

$$\varphi: G \longrightarrow G.$$

Such an isomorphism is called an *automorphism* of G. The identity map is an automorphism, of course, but there are nearly always other automorphisms as well. For example, let $G = \{1, x, x^2\}$ be a cyclic group of order 3, so that $x^3 = 1$. The transposition which interchanges x and x^2 is an automorphism of G:

$$\begin{aligned} 1 &\rightsquigarrow 1 \\ x &\rightsquigarrow x^2 \\ x^2 &\rightsquigarrow x. \end{aligned}$$

This is because x^2 is another element of order 3 in the group. If we call this element y, the cyclic subgroup $\{1, y, y^2\}$ generated by y is the whole group G, because $y^2 = x$. The automorphism compares the two realizations of G as a cyclic group.

The most important example of automorphism is conjugation: Let $b \in G$ be a fixed element. Then *conjugation by* b is the map φ from G to itself defined by

$$\varphi(x) = bxb^{-1}. \tag{3.4}$$

This is an automorphism because, first of all, it is compatible with multiplication in the group:

$$\varphi(xy) = bxyb^{-1} = bxb^{-1}byb^{-1} = \varphi(x)\varphi(y),$$

and, secondly, it is a bijective map since it has an inverse function, namely conjugation by b^{-1}. If the group is abelian, then conjugation is the identity map: $bab^{-1} = abb^{-1} = a$. But any noncommutative group has some nontrivial conjugations, and so it has nontrivial automorphisms.

The element bab^{-1} is called the *conjugate* of a by b and will appear often. Two elements a, a' of a group G are called *conjugate* if $a' = bab^{-1}$ for some $b \in G$. The conjugate behaves in much the same way as the element a itself; for example, it has the same order in the group. This follows from the fact that it is the image of a by an automorphism.

The conjugate has a useful, though trivial, interpretation. Namely, if we denote bab^{-1} by a', then

(3.5) $$ba = a'b.$$

So we can think of conjugation by b as the change in a which results when one moves b from one side to the other.

4. HOMOMORPHISMS

Let G, G' be groups. A *homomorphism* $\varphi: G \longrightarrow G'$ is any map satisfying the rule

(4.1) $$\varphi(ab) = \varphi(a)\varphi(b),$$

for all $a, b \in G$. This is the same requirement as for an isomorphism [see (3.2)]. The difference is that φ is not assumed to be bijective here.

(4.2) **Examples.** The following maps are homomorphisms:

(a) the determinant function det: $GL_n(\mathbb{R}) \longrightarrow \mathbb{R}^{\times}$;
(b) the sign of a permutation sign: $S_n \longrightarrow \{\pm 1\}$ [see Chapter 1 (4.9)];
(c) the map $\varphi: \mathbb{Z}^{+} \longrightarrow G$ defined by $\varphi(n) = a^n$, where a is a fixed element of G;
(d) the *inclusion map* $i: H \longrightarrow G$ of a subgroup H into a group G, defined by $i(x) = x$.

(4.3) **Proposition.** A group homomorphism $\varphi: G \longrightarrow G'$ carries the identity to the identity, and inverses to inverses. In other words, $\varphi(1_G) = 1_{G'}$, and $\varphi(a^{-1}) = \varphi(a)^{-1}$.

Proof. Since $1 = 1 \cdot 1$ and since φ is a homomorphism, $\varphi(1) = \varphi(1 \cdot 1) = \varphi(1)\varphi(1)$. Cancel $\varphi(1)$ from both sides by (1.12): $1 = \varphi(1)$. Next, $\varphi(a^{-1})\varphi(a) = \varphi(a^{-1}a) = \varphi(1) = 1$, and similarly $\varphi(a)\varphi(a^{-1}) = 1$. Hence $\varphi(a^{-1}) = \varphi(a)^{-1}$. □

Every group homomorphism φ determines two important subgroups: its image and its kernel. The *image* of a homomorphism $\varphi: G \longrightarrow G'$ is easy to understand. It is the image of the map

(4.4) $$\text{im } \varphi = \{x \in G' \mid x = \varphi(a) \text{ for some } a \in G\},$$

and it is a subgroup of G'. Another notation for the image is $\varphi(G)$. In Examples (4.2a,b), the image is equal to the range of the map, but in example (4.2c) it is the cyclic subgroup of G generated by a, and in Example (4.2d) it is the subgroup H.

The *kernel* of φ is more subtle. It is the set of elements of G which are mapped to the identity in G':

(4.5) $$\ker \varphi = \{a \in G \mid \varphi(a) = 1\},$$

which can also be described as the inverse image $\varphi^{-1}(1)$ of the identity element [see Appendix (1.5)]. The kernel is a subgroup of G, because if a and b are in ker φ, then $\varphi(ab) = \varphi(a)\varphi(b) = 1 \cdot 1 = 1$, hence $ab \in \ker \varphi$, and so on.

The kernel of the determinant homomorphism is the subgroup of matrices whose determinant is 1. This subgroup is called the *special linear group* and is denoted by $SL_n(\mathbb{R})$:

$$(4.6) \qquad SL_n(\mathbb{R}) = \{\text{real } n \times n \text{ matrices } A \mid \det A = 1\},$$

a subgroup of $GL_n(\mathbb{R})$. The kernel of the sign homomorphism in Example (4.2b) above is called the *alternating group* and is denoted by A_n:

$$(4.7) \qquad A_n = \{\text{even permutations}\},$$

a subgroup of S_n. The kernel of the homomorphism (4.2d) is the set of integers n such that $a^n = 1$. That this is a subgroup of $\mathbb{Z}^+$ was proved before, in (2.8).

In addition to being a subgroup, the kernel of a homomorphism has an extra property which is subtle but very important. Namely, if a is in ker φ and b is any element of the group G, then the conjugate bab^{-1} is in ker φ. For to say $a \in \ker \varphi$ means $\varphi(a) = 1$. Then

$$\varphi(bab^{-1}) = \varphi(b)\varphi(a)\varphi(b^{-1}) = \varphi(b)1\varphi(b)^{-1} = 1,$$

so $bab^{-1} \in \ker \varphi$ too.

(4.8) **Definition.** A subgroup N of a group G is called a *normal subgroup* if it has the following property: For every $a \in N$ and every $b \in G$, the conjugate bab^{-1} is in N.

As we have just seen,

(4.9) *The kernel of a homomorphism is a normal subgroup.*

Thus $SL_n(\mathbb{R})$ is a normal subgroup of $GL_n(\mathbb{R})$, and A_n is a normal subgroup of S_n.

Any subgroup of an abelian group G is normal, because when G is abelian, $bab^{-1} = a$. But subgroups need not be normal in nonabelian groups. For example, group T of invertible upper triangular matrices is not a normal subgroup of $GL_2(\mathbb{R})$. For let $A = \begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}$ and $B = \begin{bmatrix} & 1 \\ 1 & \end{bmatrix}$. Then $BAB^{-1} = \begin{bmatrix} 1 & \\ 1 & 1 \end{bmatrix}$. Here $A \in T$ and $B \in GL_2(\mathbb{R})$, but $BAB^{-1} \not\in T$.

The *center* of a group G, sometimes denoted by Z or by $Z(G)$, is the set of elements which commute with every element of G:

$$(4.10) \qquad Z = \{z \in G \mid zx = xz \text{ for all } x \in G\}.$$

The center of any group is a normal subgroup of the group. For example, it can be shown that the center of $GL_n(\mathbb{R})$ is the group of *scalar matrices*, that is, those of the form cI.

5. EQUIVALENCE RELATIONS AND PARTITIONS

A fundamental mathematical construction is to start with a set S and to form a new set by equating certain elements of S according to a given rule. For instance, we may divide the set of integers into two classes, the even integers and the odd integers. Or we may wish to view congruent triangles in the plane as equivalent geometric objects. This very general procedure arises in several ways, which we will discuss here.

Let S be a set. By a *partition* P of S, we mean a subdivision of S into nonoverlapping subsets:

(5.1) S = union of disjoint, nonempty subsets.

For example, the sets

$$\{1,3\}, \{2,5\}, \{4\}$$

form a partition of the set $\{1,2,3,4,5\}$. The two sets, of even integers and of odd integers, form a partition of the set $\mathbb{Z}$ of all integers.

An *equivalence relation* on S is a relation which holds between certain elements of S. We often write it as $a \sim b$ and speak of it as *equivalence* of a and b.

(5.2) An equivalence relation is required to be:

(i) *transitive:* If $a \sim b$ and $b \sim c$, then $a \sim c$;

(ii) *symmetric:* If $a \sim b$, then $b \sim a$;

(iii) *reflexive:* $a \sim a$ for all $a \in S$.

Congruence of triangles is an example of an equivalence relation on the set S of triangles in the plane.

Formally, a relation on S is the same thing as a subset R of the set $S \times S$ of pairs of elements; namely, the subset R consists of pairs (a,b) such that $a \sim b$. In terms of this subset, we can write the axioms for an equivalence relation as follows: (i) if $(a,b) \in R$ and $(b,c) \in R$, then $(a,c) \in R$; (ii) if $(a,b) \in R$, then $(b,a) \in R$; and (iii) $(a,a) \in R$ for all a.

The notions of a partition of S and an equivalence relation on S are logically equivalent, though in practice one is often presented with just one of the two. Given a partition P on S, we can define an equivalence relation R by the rule $a \sim b$ if a and b lie in the same subset of the partition. Axioms (5.2) are obviously satisfied. Conversely, given an equivalence relation R, we can define a partition P this way: The subset containing a is the set of all elements b such that $a \sim b$. This subset is called the *equivalence class* of a, and S is partitioned into equivalence classes.

Let us check that the equivalence classes partition the set S. Call C_a the equivalence class of an element $a \in S$. So C_a consists of the elements b such that $a \sim b$:

(5.3) $$C_a = \{b \in S \mid a \sim b\}.$$

The reflexive axiom tells us that $a \in C_a$. Therefore the classes C_a are nonempty, and since a can be any element, the classes cover S. The remaining property of a partition which must be verified is that equivalence classes do not overlap. It is easy to become confused here, because if $a \sim b$ then by definition $b \in C_a$. But $b \in C_b$ too. Doesn't this show that C_a and C_b overlap? We must remember that the symbol C_a is our notation for a subset of S defined in a certain way. The partition consists of the subsets, not of the notations. It is true that C_a and C_b have the element b in common, but that is all right because these are two notations for the same set. We will show the following:

(5.4) *Suppose that C_a and C_b have an element d in common. Then $C_a = C_b$.*

Let us first show that if $a \sim b$ then $C_a = C_b$. To do so, let x be an arbitrary element of C_b. Then $b \sim x$. Since $a \sim b$, transitivity shows that $a \sim x$, hence that $x \in C_a$. Therefore $C_b \subset C_a$. The opposite inclusion follows from interchanging the roles of a and b. To prove (5.4), suppose that d is in C_a and in C_b; then $a \sim d$ and $b \sim d$. Then by what has been shown, $C_a = C_d = C_b$, as required. □

Suppose that an equivalence relation or a partition is given on a set S. Then we may construct a new set $\overline{S}$ whose elements are the equivalence classes or the subsets making up the partition. To simplify notation, the equivalence class of a, or the subset of the partition containing a, is often denoted by $\overline{a}$. Thus $\overline{a}$ is an element of $\overline{S}$.

Notice that there is a natural surjective map

$$S \longrightarrow \overline{S}, \text{ which sends} \qquad (5.5)$$
$$a \rightsquigarrow \overline{a}.$$

In our original example of the partition of $S = \mathbb{Z}$, the set $\overline{S}$ contains the two elements $(Even)$, (Odd), where the symbol $\overline{(Even)}$ represents the set of even integers and (Odd) the set of odd integers. And $\overline{0} = \overline{2} = \overline{4}$ and so on. So we can denote the set $(Even)$ by any one of these symbols. The map

$$\mathbb{Z} \longrightarrow \{(Even), (Odd)\} \qquad (5.6)$$

is the obvious one.

There are two ways to think of this construction. We can imagine putting the elements of S into separate piles, one for each subset of the partition, and then regarding the piles as the elements of a new set $\overline{S}$. The map $S \longrightarrow \overline{S}$ associates each element with its pile. Or we can think of changing what we mean by equality among elements of S, interpreting $a \sim b$ to mean $a = b$ in $\overline{S}$. With this way of looking at it, the elements in the two sets S and $\overline{S}$ correspond, but in $\overline{S}$ more of them are equal to each other. It seems to me that this is the way we treat congruent triangles in school. The bar notation (5.5) is well suited to this intuitive picture. We can work with the same symbols as in S, but with bars over them to remind us of the new rule:

$$\overline{a} = \overline{b} \text{ means } a \sim b. \qquad (5.7)$$

This notation is often very convenient.

A disadvantage of the bar notation is that many symbols represent the same element of $\overline{S}$. Sometimes this disadvantage can be overcome by choosing once and for all a particular element, or a *representative,* in each equivalence class. For example, it is customary to represent $(Even)$ by $\overline{0}$ and (Odd) by $\overline{1}$:

$$\{(Even), (Odd)\} = \{\overline{0}, \overline{1}\}. \tag{5.8}$$

Though the pile picture is more immediate, the second way of viewing $\overline{S}$ is often the better one, because operations on the piles are clumsy to visualize, whereas the bar notation is well suited to algebraic manipulation.

Any map of sets $\varphi: S \longrightarrow T$ defines an equivalence relation on the domain S, namely the relation given by the rule $a \sim b$ if $\varphi(a) = \varphi(b)$. We will refer to this as *the equivalence relation determined by the map.* The corresponding partition is made up of the nonempty inverse images of the elements of T. By definition, the *inverse image* of an element $t \in T$ is the subset of S consisting of all elements s such that $\varphi(s) = t$. It is denoted symbolically as

$$\varphi^{-1}(t) = \{s \in S \mid \varphi(s) = t\}. \tag{5.9}$$

Thus $\varphi^{-1}(t)$ is a subset of the domain S, determined by the element $t \in T$. (This is symbolic notation. Please remember that φ^{-1} is usually not a function.) The inverse images may also be called the *fibres* of the map φ. The fibres $\varphi^{-1}(t)$ which are *nonempty,* which means t is in the image of φ, form a partition of S. Here the set $\overline{S}$ of equivalence classes, which is the set of nonempty fibres, has another incarnation, as the image im φ of the map. Namely, there is a bijective map

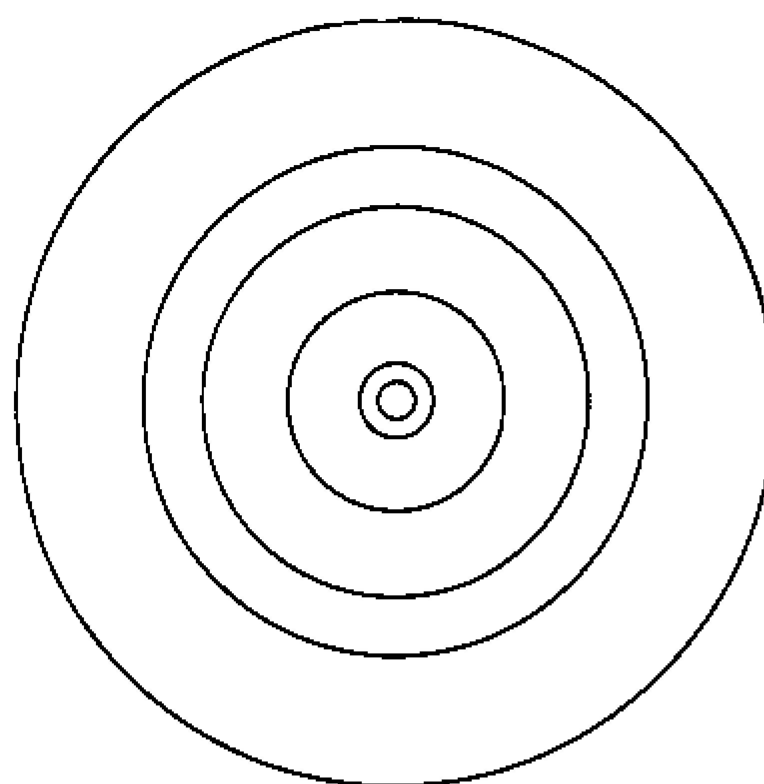
$$\overline{\varphi}: \overline{S} \longrightarrow \text{im } \varphi, \tag{5.10}$$

the map which sends an element $\overline{s}$ of $\overline{S}$ to $\varphi(s)$.

We now go back to group homomorphisms. Let $\varphi: G \longrightarrow G'$ be a homomorphism, and let us analyze the equivalence relation on G which is associated to the map φ or, equivalently, the fibres of the homomorphism. This relation is usually denoted by $\equiv$, rather than by $\sim$, and is referred to as *congruence:*

$$a \equiv b \quad \text{if} \quad \varphi(a) = \varphi(b). \tag{5.11}$$

For example, let $\varphi: \mathbb{C}^\times \longrightarrow \mathbb{R}^\times$ be the absolute value homomorphism defined by $\varphi(a) = |a|$. The induced equivalence relation is $a \equiv b$ if $|a| = |b|$. The fibres of this map are the concentric circles about 0. They are in bijective correspondence with elements of im φ, the set of positive reals.



(5.12) **Figure.** Fibres of the absolute value map $\mathbb{C}^\times \longrightarrow \mathbb{R}^\times$.

The relation (5.11) can be rewritten in a number of ways, of which the following will be the most important for us:

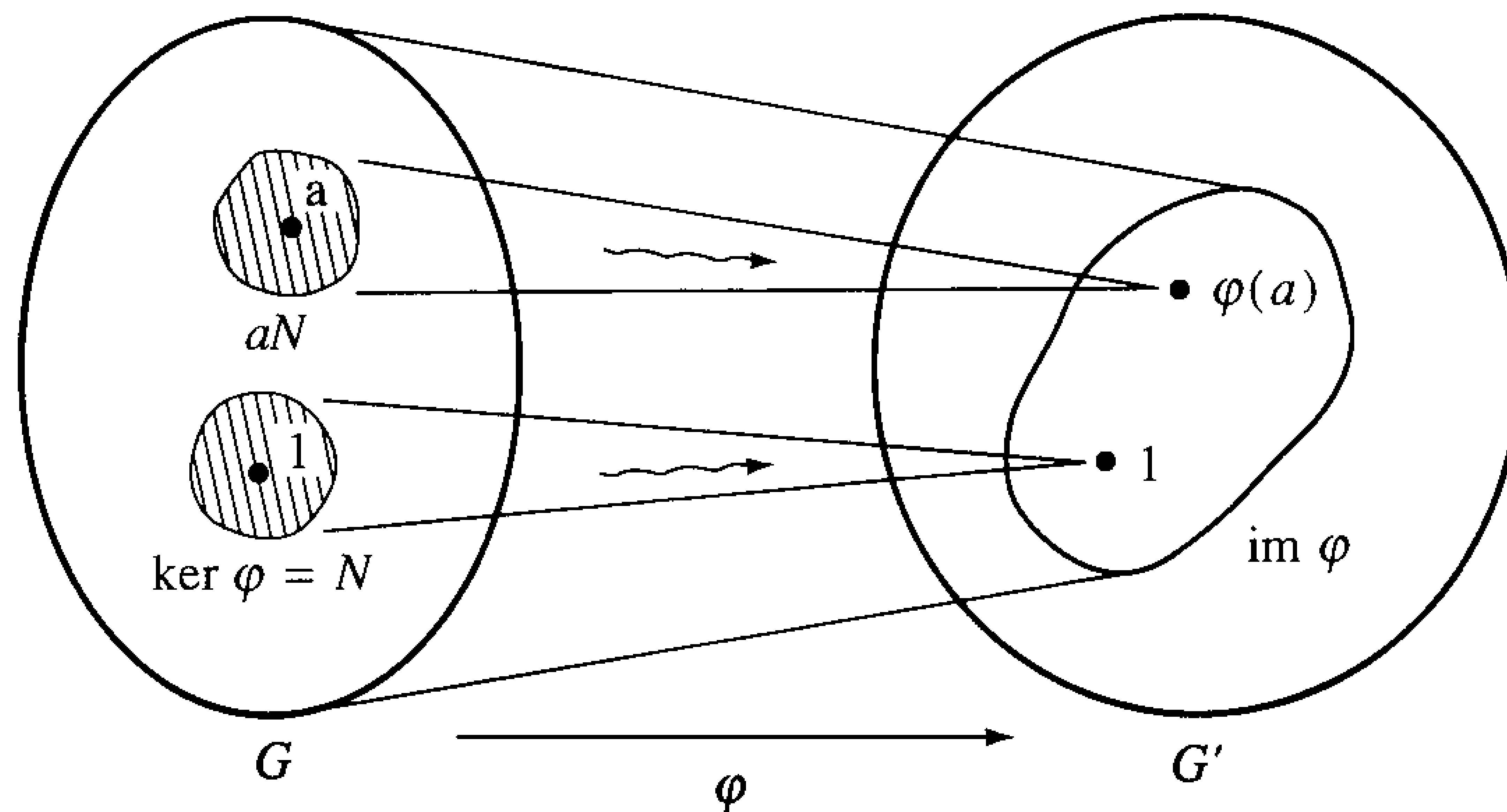
(5.13) Proposition. Let $\varphi: G \longrightarrow G'$ be a group homomorphism with kernel N, and let a, b be elements of G. Then $\varphi(a) = \varphi(b)$ if and only if $b = an$ for some element $n \in N$, or equivalently, if $a^{-1}b \in N$.

Proof. Suppose that $\varphi(a) = \varphi(b)$. Then $\varphi(a)^{-1}\varphi(b) = 1$, and since φ is a homomorphism we can use (4.1) and (4.3) to rewrite this equality as $\varphi(a^{-1}b) = 1$. Now by definition, the kernel N is the set of all elements $x \in G$ such that $\varphi(x) = 1$. Thus $a^{-1}b \in N$, or $a^{-1}b = n$ for some $n \in N$. Hence $b = an$, as required. Conversely, if $b = an$ and $n \in N$, then $\varphi(b) = \varphi(a)\varphi(n) = \varphi(a)1 = \varphi(a)$. □

The set of elements of the form an is denoted by aN and is called a *coset* of N in G:

(5.14) $$aN = \{g \in G \mid g = an \text{ for some } n \in N\}.$$

So the coset aN is the set of all group elements b which are congruent to a. The congruence relation $a \equiv b$ partitions the group G into *congruence classes*, the cosets aN. They are the fibres of the map φ. In particular, the circles about the origin depicted in (5.12) are cosets of the absolute value homomorphism.



(5.15) **Figure.** A schematic diagram of a group homomorphism.

An important case to look at is when the kernel is the trivial subgroup. In that case (5.13) reads as follows:

(5.16) Corollary. A group homomorphism $\varphi: G \longrightarrow G'$ is injective if and only if its kernel is the trivial subgroup $\{1\}$. □

This gives us a way to verify that a homomorphism is an isomorphism. To do so, we check that $\ker \varphi = \{1\}$, so that φ is injective, and also that $\operatorname{im} \varphi = G'$, that is, that φ is surjective.

6. COSETS

One can define cosets for any subgroup H of a group G, not only for the kernel of a homomorphism. A *left coset* is a subset of the form

(6.1) $$aH = \{ah \mid h \in H\}.$$

Note that the subgroup H is itself a coset, because $H = 1H$.

The cosets are equivalence classes for the *congruence* relation

(6.2) $$a \equiv b \text{ if } b = ah, \text{ for some } h \in H.$$

Let us verify that congruence is an equivalence relation. *Transitivity:* Suppose that $a \equiv b$ and $b \equiv c$. This means that $b = ah$ and $c = bh'$ for some $h,h' \in H$. Therefore $c = ahh'$. Since H is a subgroup, $hh' \in H$. Thus $a \equiv c$. *Symmetry:* Suppose $a \equiv b$, so that $b = ah$. Then $a = bh^{-1}$ and $h^{-1} \in H$, and so $b \equiv a$. *Reflexivity:* $a = a1$ and $1 \in H$, so $a \equiv a$. Note that we have made use of all the defining properties of a subgroup.

Since equivalence classes form a partition, we find the following:

(6.3) **Corollary.** The left cosets of a subgroup partition the group. □

(6.4) **Note.** The notation aH defines a certain subset of G. As with any equivalence relation, different notations may represent the same subset. In fact, we know that aH is the unique coset containing a, and so

(6.5) $$aH = bH \textit{ if and only if } a \equiv b.$$

The corollary just restates (5.4):

(6.6) *If aH and bH have an element in common, then they are equal.*

For example, let G be the symmetric group S_3, with the presentation given in (1.18): $G = \{1, x, x^2, y, xy, x^2y\}$. The element xy has order 2, and so it generates a cyclic subgroup $H = \{1, xy\}$ of order 2. The left cosets of H in G are the three sets

(6.7) $$\{1, xy\} = H = xyH, \quad \{x, x^2y\} = xH = x^2yH, \quad \{x^2y\} = x^2H = yH.$$

Notice that they do partition the group.

The number of left cosets of a subgroup is called the *index* of H in G and is denoted by

(6.8) $$[G : H].$$

Thus in our example the index is 3. Of course if G contains infinitely many elements, the index may be infinite too.

Note that there is a bijective map from the subgroup H to the coset aH, sending $h \rightsquigarrow ah$. (Why is this a bijective map?) Thus

(6.9) *Each coset aH has the same number of elements as H does.*

Since G is the union of the cosets of H and since these cosets do not overlap, we obtain the important *Counting Formula*

(6.10) $$|G| = |H|[G : H],$$

where $|G|$ denotes the order of the group, as in (2.10), and where the equality has the obvious meaning if some terms are infinite. In our example (6.7), this formula reads $6 = 2 \cdot 3$.

The fact that the two terms on the right side of equation (6.10) must divide the left side is very important. Here is one of these conclusions, stated formally:

(6.11) **Corollary.** *Lagrange's Theorem:* Let G be a finite group, and let H be a subgroup of G. The order of H divides the order of G. □

In Section 2 we defined the order of an element $a \in G$ to be the order of the cyclic subgroup generated by a. Hence Lagrange's Theorem implies the following:

(6.12) *The order of an element divides the order of the group.*

This fact has a remarkable consequence:

(6.13) **Corollary.** Suppose that a group G has p elements and that p is a prime integer. Let $a \in G$ be any element, not the identity. Then G is the cyclic group $\{1, a, \dots, a^{p-1}\}$ generated by a.

For, since $a \neq 1$, the order of a is greater than 1, and it divides $|G| = p$. Hence it is equal to p. Since G has order p, $\{1, a, \dots, a^{p-1}\}$ is the whole group. □

Thus we have classified all groups of prime order p. They form one isomorphism class, the class of a cyclic group of order p.

The Counting Formula can also be applied when a homomorphism is given. Let $\varphi: G \longrightarrow G'$ be a homomorphism. As we saw in (5.13), the left cosets of $\ker \varphi$ are the fibres of the map φ. They are in bijective correspondence with the elements in the image.

(6.14) $$[G : \ker \varphi] = |\operatorname{im} \varphi|.$$

Thus (6.10) implies the following:

(6.15) **Corollary.** Let $\varphi: G \longrightarrow G'$ be a homomorphism of finite groups. Then

$$|G| = |\ker \varphi| \cdot |\operatorname{im} \varphi|.$$

Thus $|\ker \varphi|$ divides $|G|$, and $|\operatorname{im} \varphi|$ divides both $|G|$ and $|G'|$.

Proof. The formula is obtained by combining (6.10) and (6.14), and it implies that $|\ker \varphi|$ and $|\operatorname{im} \varphi|$ divide $|G|$. Since $\operatorname{im} \varphi$ is a subgroup of G', $|\operatorname{im} \varphi|$ divides $|G'|$ as well. □

Let us go back for a moment to the definition of cosets. We made the decision to work with left cosets aH. One can also define right cosets of a subgroup H and repeat the above discussion for them. The *right cosets* of a subgroup H are the sets

$$Ha = \{ha \mid h \in H\}, \tag{6.16}$$

which are equivalence classes for the relation (*right congruence*)

$$a \equiv b \text{ if } b = ha, \text{ for some } h \in H.$$

Right cosets need not be the same as left cosets. For instance, the right cosets of the subgroup $\{1, xy\}$ of S_3 are

$$\{1, xy\} = H = Hxy, \quad \{x, y\} = Hx = Hy, \quad \{x^2, x^2y\} = Hx^2 = Hx^2y. \tag{6.17}$$

This partition of S_3 is not the same as the partition (6.7) into left cosets.

However, if N is a normal subgroup, then right and left cosets agree.

(6.18) **Proposition.** A subgroup H of a group G is normal if and only if every left coset is also a right coset. If H is normal, then $aH = Ha$ for every $a \in G$.

Proof. Suppose that H is normal. For any $h \in H$ and any $a \in G$,

$$ah = (aha^{-1})a.$$

Since H is a normal subgroup, the conjugate element $k = aha^{-1}$ is in H. Thus the element $ah = ka$ is in aH and also in Ha. This shows that $aH \subset Ha$. Similarly, $aH \supset Ha$, and so these two cosets are equal. Conversely, suppose that H is not normal. Then there are elements $h \in H$ and $a \in G$ so that aha^{-1} is not in H. Then ah is in the left coset aH but not in the right coset Ha. If it were, say $ah = h'a$ for some $h' \in H$, then we would have $aha^{-1} = h' \in H$, contrary to our hypothesis. On the other hand, aH and Ha do have an element in common, namely the element a. So aH can't be in some other right coset. This shows that the partition into left cosets is not the same as the partition into right cosets. □

7. RESTRICTION OF A HOMOMORPHISM TO A SUBGROUP

The usual way to get an understanding of a complicated group is to study some less complicated subgroups. If it made sense to single out one method in group theory as the most important, this would be it. For example, the general linear group GL_2 is much more complicated than the group of invertible upper triangular matrices. We expect to answer any question about upper triangular matrices which comes up. And by taking products of upper and lower triangular matrices, we can cover most of the group GL_2. Of course, the trick is to get back information about a group from an understanding of its subgroups. We don't have general rules about how this should be done. But whenever a new construction with groups is made, we should study its effect on subgroups. This is what is meant by *restriction to a subgroup*. We will do this for subgroups and homomorphisms in this section.

Let H be a subgroup of a group G. Let us first consider the case that a second subgroup K is given. The restriction of K to H is the intersection $K \cap H$. The following proposition is a simple exercise.

(7.1) **Proposition.** The intersection $K \cap H$ of two subgroups is a subgroup of H. If K is a normal subgroup of G, then $K \cap H$ is a normal subgroup of H. □

There is not very much more to be said here, but if G is a finite group, we may be able to apply the Counting Formula (6.10), especially Lagrange's Theorem, to get information about the intersection. Namely, $K \cap H$ is a subgroup of H and also a subgroup of K. So its order divides both of the orders $|H|$ and $|K|$. If $|H|$ and $|K|$ have no common factor, we can conclude that $K \cap H = \{1\}$.

Now suppose that a homomorphism $\varphi: G \longrightarrow G'$ is given and that H is a subgroup of G as before. Then we may *restrict* φ *to* H, obtaining a homomorphism

$$\varphi|_H: H \longrightarrow G'. \tag{7.2}$$

This means that we take the same map φ but restrict its domain to H. In other words, $\varphi|_H(h) = \varphi(h)$ for all $h \in H$. The restriction is a homomorphism because φ is one.

The kernel of $\varphi|_H$ is the intersection of ker φ with H :

$$\ker \varphi|_H = (\ker \varphi) \cap H. \tag{7.3}$$

This is clear from the definition of kernel: $\varphi(h) = 1$ if and only if $h \in \ker \varphi$.

Again, the Counting Formula may help to describe this restriction. For, the image of $\varphi|_H$ is $\varphi(H)$. According to Corollary (6.15), $|\varphi(H)|$ divides both $|H|$ and $|G'|$. So if $|H|$ and $|G'|$ have no common factor, $\varphi(H) = \{1\}$. Then we can conclude that $H \subset \ker \varphi$.

For example, the sign of a permutation is described by a homomorphism (4.2b), $S_n \longrightarrow \{\pm 1\}$. The range of this homomorphism has order 2, and its kernel is the alternating group. If a subgroup H of S_n has odd order, then the restriction of this homomorphism to H is trivial, which means that H is contained in the alternating group, that is, H consists of even permutations. This will be so when H is the cyclic subgroup generated by a permutation p whose order in the group is odd. It follows that every permutation of odd order is an even permutation. On the other hand, we can not make any conclusion about permutations of even order. They may be odd or even.

When a homomorphism $\varphi: G \longrightarrow G'$ and a subgroup H' of G' are given, we may also restrict φ to H'. Here we must cut down the domain G of φ suitably, in order to get a map to H'. The natural thing to do is to cut down the domain as little as possible by taking the entire inverse image of H':

(7.4) **Proposition.** Let $\varphi: G \longrightarrow G'$ be a homomorphism, and let H' be a subgroup of G'. Denote the inverse image $\varphi^{-1}(H') = \{x \in G \mid \varphi(x) \in H'\}$ by $\tilde{H}$. Then

(a) $\tilde{H}$ is a subgroup of G.
(b) If H' is a normal subgroup of G', then $\tilde{H}$ is a normal subgroup of G.
(c) $\tilde{H}$ contains $\ker \varphi$.
(d) The restriction of φ to $\tilde{H}$ defines a homomorphism $\tilde{H} \longrightarrow H'$, whose kernel is $\ker \varphi$.

For example, consider the determinant homomorphism $\det: GL_n(\mathbb{R}) \longrightarrow \mathbb{R}^\times$. The set P of positive real numbers is a subgroup of $\mathbb{R}^\times$, and its inverse image is the set of invertible $n \times n$ matrices with positive determinant, which is a normal subgroup of $GL_n(\mathbb{R})$.

Proof of Proposition (7.4). This proof is also a simple exercise, but we must keep in mind that φ^{-1} is not a map. By definition, $\tilde{H}$ is the set of elements $x \in G$ such that $\varphi(x) \in H'$. We verify the conditions for a subgroup. *Identity:* $1 \in \tilde{H}$ because $\varphi(1) = 1 \in H'$. *Closure:* Suppose that $x, y \in \tilde{H}$. This means that $\varphi(x)$ and $\varphi(y)$ are in H'. Since H' is a subgroup, $\varphi(x)\varphi(y) \in H'$. Since φ is a homomorphism, $\varphi(x)\varphi(y) = \varphi(xy) \in H'$. Therefore $xy \in \tilde{H}$. *Inverses:* Suppose $x \in \tilde{H}$, so that $\varphi(x) \in H'$; then $\varphi(x)^{-1} \in H'$ because H' is a subgroup. Since φ is a homomorphism, $\varphi(x)^{-1} = \varphi(x^{-1})$. Thus $x^{-1} \in \tilde{H}$.

Suppose that H' is a normal subgroup, and let $x \in \tilde{H}$ and $g \in G$. Then $\varphi(gxg^{-1}) = \varphi(g)\varphi(x)\varphi(g)^{-1}$, and $\varphi(x) \in H'$. Therefore $\varphi(gxg^{-1}) \in H'$, and this shows that $gxg^{-1} \in \tilde{H}$. Next, $\tilde{H}$ contains $\ker \varphi$ because if $x \in \ker \varphi$ then $\varphi(x) = 1$, and $1 \in H'$. So $x \in \varphi^{-1}(H')$. The last assertion should be clear. □

8. PRODUCTS OF GROUPS

Let G, G' be two groups. The product set $G \times G'$ can be made into a group by component-wise multiplication. That is, we define multiplication of pairs by the rule

(8.1) $$(a, a'), (b, b') \rightsquigarrow (ab, a'b'),$$

for $a, b \in G$ and $a', b' \in G'$. The pair $(1, 1)$ is an identity, and $(a, a')^{-1} = (a^{-1}, a'^{-1})$. The associative law in $G \times G'$ follows from the fact that it holds in G and in G'. The group thus obtained is called the *product* of G and G' and is denoted by $G \times G'$. Its order is the product of the orders of G and G'.

The product group is related to the two factors G, G' in a simple way, which we can sum up in terms of some homomorphisms

(8.2) $$\begin{array}{ccccc} G & & & & G \\ & \overset{i}{\searrow} & & \overset{p}{\nearrow} & \\ & & G \times G' & & \\ & \underset{i'}{\nearrow} & & \underset{p'}{\searrow} & \\ G' & & & & G' \end{array},$$

defined by

$$i(x) = (x, 1), \quad i'(x') = (1, x'),$$

$$p(x, x') = x, \quad p'(x, x') = x'.$$

The maps i, i' are injective and may be used to identify G, G' with the subgroups $G \times 1$, $1 \times G'$ of $G \times G'$. The maps p, p' are surjective, $\ker p = 1 \times G'$, and $\ker p' = G \times 1$. These maps are called the *projections*. Being kernels, $G \times 1$ and $1 \times G'$ are *normal* subgroups of $G \times G'$.

(8.3) **Proposition.** *The mapping property of products:* Let H be any group. The homomorphisms $\Phi: H \longrightarrow G \times G'$ are in bijective correspondence with pairs (φ, φ') of homomorphisms

$$\varphi: H \longrightarrow G, \qquad \varphi': H \longrightarrow G'.$$

The kernel of Φ is the intersection $(\ker \varphi) \cap (\ker \varphi')$.

Proof. Given a pair (φ, φ') of homomorphisms, we define the corresponding homomorphism

$$\Phi: H \longrightarrow G \times G'$$

by the rule $\Phi(h) = (\varphi(h), \varphi'(h))$. This is easily seen to be a homomorphism. Conversely, given Φ, we obtain φ and φ' by composition with the projections, as

$$\varphi = p\Phi, \qquad \varphi' = p'\Phi.$$

Obviously, $\Phi(h) = (1, 1)$ if and only if $\varphi(h) = 1$ and $\varphi'(h) = 1$, which shows that $\ker \Phi = (\ker \varphi) \cap (\ker \varphi')$. □

It is clearly desirable to compose a given group G as a product, meaning to find two groups H and H' such that G is isomorphic to the product $H \times H'$. For the groups H, H' will be smaller and therefore simpler, and the relation between $H \times H'$ and its factors is easily understood. Unfortunately, it is quite rare that a given group is a product, but it does happen occasionally.

For example, it is rather surprising that a cyclic group of order 6 can be decomposed: A cyclic group C_6 of order 6 is isomorphic to the product $C_2 \times C_3$ of cyclic groups of orders 2 and 3. This can be shown using the mapping property just discussed. Say that $C_6 = \{1, x, x^2, \ldots, x^5\}$, $C_2 = \{1, y\}$, $C_3 = \{1, z, z^2\}$. The rule

$$\varphi: C_6 \longrightarrow C_2 \times C_3$$

defined by $\varphi(x^i) = (y^i, z^i)$ is a homomorphism, and its kernel is the set of elements x^i such that $y^i = 1$ and $z^i = 1$. Now $y^i = 1$ if and only if i is divisible by 2, while $z^i = 1$ if and only if i is divisible by 3. There is no integer between 1 and 5 which is divisible by both 2 and 3. Therefore $\ker \varphi = \{1\}$, and φ is injective. Since both groups have order 6, φ is bijective and hence is an isomorphism. □

The same argument works for a *cyclic* group of order rs, whenever the two integers r and s have no common factor.

(8.4) **Proposition.** Let r, s be integers with no common factor. A *cyclic* group of order rs is isomorphic to the product of a cyclic group of order r and a cyclic group of order s. □

On the other hand, a cyclic group of order 4 is *not* isomorphic to a product of two cyclic groups of order 2. For it is easily seen that every element of $C_2 \times C_2$ has order 1 or 2, whereas a cyclic group of order 4 contains two elements of order 4. And, the proposition makes no assertions about a group which is not cyclic.

Let A and B be subsets of a group G. Then we denote the set of products of elements of A and B by

$$AB = \{x \in G \mid x = ab \text{ for some } a \in A \text{ and } b \in B\}. \tag{8.5}$$

The next proposition characterizes product groups.

(8.6) **Proposition.** Let H and K be subgroups of a group G.

(a) If $H \cap K = \{1\}$, the product map $p: H \times K \longrightarrow G$ defined by $p(h,k) = hk$ is injective. Its image is the subset HK.
(b) If either H or K is a normal subgroup of G, then the product sets HK and KH are equal, and HK is a subgroup of G.
(c) If H and K are normal, $H \cap K = \{1\}$, and $HK = G$, then G is isomorphic to the product group $H \times K$.

Proof. (a) Let (h_1, k_1), (h_2, k_2) be elements of $H \times K$ such that $h_1k_1 = h_2k_2$. Multiplying both sides of this equation on the left by h_1^{-1} and on the right by k_2^{-1}, we find $k_1k_2^{-1} = h_1^{-1}h_2$. Since $H \cap K = \{1\}$, $k_1k_2^{-1} = h_1^{-1}h_2 = 1$, hence $h_1 = h_2$ and $k_1 = k_2$. This shows that p is injective.

(b) Suppose that H is a normal subgroup of G, and let $h \in H$ and $k \in K$. Note that $kh = (khk^{-1})k$. Since H is normal, $khk^{-1} \in H$. Therefore $kh \in HK$, which shows that $KH \subset HK$. The proof of the other inclusion is similar. The fact that HK is a subgroup now follows easily. For closure under multiplication, note that in a product $(hk)(h'k') = h(kh')k'$, the middle term kh' is in $KH = HK$, say $kh' = h''k''$. Then $hkh'k' = (hh'')(k''k') \in HK$. Closure under inverses is similar: $(hk)^{-1} = k^{-1}h^{-1} \in KH = HK$. And of course, $1 = 1 \cdot 1 \in HK$. Thus HK is a subgroup. The proof is similar in the case that K is normal.

(c) Assume that both subgroups are normal and that $H \cap K = \{1\}$. Consider the product $(hkh^{-1})k^{-1} = h(kh^{-1}k^{-1})$. Since K is a normal subgroup, the left side is in K. Since H is normal, the right side is in H. Thus this product is the intersection $H \cap K$, i.e., $hkh^{-1}k^{-1} = 1$. Therefore $hk = kh$. This being known, the fact that p is a homomorphism follows directly: In the group $H \times K$, the product rule is $(h_1, k_1)(h_2, k_2) = (h_1h_2, k_1k_2)$, and this element corresponds to $h_1h_2k_1k_2$ in G, while

in G the products h_1k_1 and h_2k_2 multiply as $h_1k_1h_2k_2$. Since $h_2k_1 = k_1h_2$, the products are equal. Part (a) shows that p is injective, and the assumption that $HK = G$ shows that p is surjective. □

It is important to note that the product map $p: H \times K \longrightarrow G$ will not be a group homomorphism unless the two subgroups commute with each other.

9. MODULAR ARITHMETIC

In this section we discuss Gauss's definition of congruence of integers, which is one of the most important concepts in number theory. We work with a fixed, but arbitrary, positive integer n throughout this section.

Two integers a,b are said to be *congruent modulo n*, written

(9.1) $$a \equiv b \text{ (modulo } n),$$

if n divides $b - a$, or if $b = a + nk$ for some integer k. It is easy to check that this is an equivalence relation. So we may consider the equivalence classes, called *congruence classes modulo n* or *residue classes modulo n*, defined by this relation, as in Section 5. Let us denote the congruence class of an integer a by the symbol $\overline{a}$. It is the set of integers

(9.2) $$\overline{a} = \{\ldots, a - 2n, a - n, a, a + n, a + 2n, \ldots\}.$$

If a and b are integers, the equation $\overline{a} = \overline{b}$ means that n divides $b - a$.

The congruence class of 0 is the subgroup

$$\overline{0} = n\mathbb{Z} = \{\ldots, -n, 0, n, 2n, \ldots\}$$

of the additive group $\mathbb{Z}^+$ consisting of all multiples of n. The other congruence classes are the cosets of this subgroup. Unfortunately, we have a slight notational problem here, because the notation $n\mathbb{Z}$ is like the one we use for a coset. But $n\mathbb{Z}$ is not a coset; it is a subgroup of $\mathbb{Z}^+$. The notation for a coset of a subgroup H analogous to (6.1), but using additive notation for the law of composition, is

$$a + H = \{a + h \mid h \in H\}.$$

In order to avoid writing a coset as $a + n\mathbb{Z}$, let us denote the subgroup $n\mathbb{Z}$ by H. Then the cosets of H are the sets

(9.3) $$a + H = \{a + nk \mid k \in \mathbb{Z}\}.$$

They are the congruence classes $\overline{a} = a + H$.

The n integers $0,1,\ldots,n - 1$ form a natural set of representative elements for the congruence classes:

(9.4) **Proposition.** There are n congruence classes modulo n, namely

$$\overline{0}, \overline{1}, \ldots, \overline{n - 1}.$$

Or, the index $[\mathbb{Z} : n\mathbb{Z}]$ of the subgroup $n\mathbb{Z}$ in $\mathbb{Z}$ is n.

Proof. Let a be an arbitrary integer. Then we may use division with remainder to write

$$a = nq + r,$$

where q, r are integers and where the remainder r is in the range $0 \leq r < n$. Then a is congruent to the remainder: $a \equiv r$ (modulo n). Thus $\overline{a} = \overline{r}$. This shows that $\overline{a}$ is one of the congruence classes listed in the proposition. On the other hand, if a and b are distinct integers less than n, say $a \leq b$, then $b - a$ is less than n and different from zero, so n does not divide $b - a$. Thus $a \not\equiv b$ (modulo n), which means that $\overline{a} \neq \overline{b}$. Therefore the n classes $\overline{0}, \overline{1}, \ldots, \overline{n-1}$ are distinct. □

The main point about congruence classes is that addition and multiplication of integers preserve congruences modulo n, and therefore these laws can be used to define addition and multiplication of congruence classes. This is expressed by saying that the set of congruence classes forms a *ring*. We will study rings in Chapter 10.

Let $\overline{a}$ and $\overline{b}$ be congruence classes represented by integers a and b. Their *sum* is defined to be the congruence class of $a + b$, and their *product* is defined to be the class of ab. In other words, we define

$$\overline{a} + \overline{b} = \overline{a + b} \quad \text{and} \quad \overline{a}\overline{b} = \overline{ab}. \tag{9.5}$$

This definition needs some justification, because the same congruence class $\overline{a}$ can be represented by many different integers. Any integer a' congruent to a modulo n represents the same class. So it had better be true that if $a' \equiv a$ and $b' \equiv b$, then $a' + b' \equiv a + b$ and $a'b' \equiv ab$. Fortunately, this is so.

(9.6) **Lemma.** If $a' \equiv a$ and $b' \equiv b$ (modulo n), then $a' + b' \equiv a + b$ (modulo n) and $a'b' \equiv ab$ (modulo n).

Proof. Assume that $a' \equiv a$ and $b' \equiv b$, so that $a' = a + nr$ and $b' = b + ns$ for some integers r, s. Then $a' + b' = a + b + n(r + s)$, which shows that $a' + b' \equiv a + b$. Similarly, $a'b' = (a + nr)(b + ns) = ab + n(as + rb + nrs)$, which shows that $a'b' \equiv ab$, as required. □

The associative, commutative, and distributive laws hold for the laws of composition (9.5) because they hold for addition and multiplication of integers. For example, the formal verification of the distributive law is as follows:

$$\begin{aligned}
\overline{a}(\overline{b} + \overline{c}) = \overline{a}(\overline{b + c}) = \overline{a(b + c)} \quad & \textit{(definition of + and × for congruence classes)} \\
= \overline{ab + ac} \quad & \textit{(distributive law in the integers)} \\
= \overline{ab} + \overline{ac} = \overline{a}\overline{b} + \overline{a}\overline{c} \quad & \textit{(definition of + and × for congruence classes).}
\end{aligned}$$

The set of congruence classes modulo n is usually denoted by

$$\mathbb{Z}/n\mathbb{Z}. \tag{9.7}$$

Computation of addition, subtraction, and multiplication in $\mathbb{Z}/n\mathbb{Z}$ can be made ex-

plicitly by working with integers and taking remainders on division by n. That is what the formulas (9.5) mean. They tell us that the map

(9.8) $$\mathbb{Z} \longrightarrow \mathbb{Z}/n\mathbb{Z}$$

sending an integer a to its congruence class $\overline{a}$ is compatible with addition and multiplication. Therefore computations can be made in the integers and then carried over to $\mathbb{Z}/n\mathbb{Z}$ at the end. However, doing this is not efficient, because computations are simpler if the numbers are kept small. We can keep them small by computing the remainder after some part of a computation has been made.

Thus if $n = 13$, so that

$$\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \overline{1}, \overline{2}, \dots, \overline{12}\},$$

then

$$(\overline{7} + \overline{9})(\overline{11} + \overline{6})$$

can be computed as $\overline{7} + \overline{9} = \overline{3}$, $\overline{11} + \overline{6} = \overline{4}$, $\overline{3} \cdot \overline{4} = \overline{12}$.

The bars over the numbers are a nuisance, so they are often left off. One just has to remember the following rule:

(9.9) *To say $a = b$ in $\mathbb{Z}/n\mathbb{Z}$ means $a \equiv b$* (modulo n).

10. QUOTIENT GROUPS

We saw in the last section that the congruence classes of integers modulo n are the cosets of the subgroup $n\mathbb{Z}$ of $\mathbb{Z}^+$. So addition of congruence classes gives us a law of composition on the set of these cosets. In this section we will show that a law of composition can be defined on the cosets of a normal subgroup N of any group G. We will show how to make the set of cosets into a group, called a *quotient group*.

Addition of angles is a familiar example of the quotient construction. Every real number represents an angle, and two real numbers represent the same angle if they differ by an integer multiple of 2π. This is very familiar. The point of the example is that addition of angles is defined in terms of addition of real numbers. The group of angles is a quotient group, in which $G = \mathbb{R}^+$ and N is the subgroup of integer multiples of 2π.

We recall a notation introduced in Section 8: If A and B are subsets of a group G, then

$$AB = \{ab \mid a \in A, b \in B\}.$$

We will call this the *product* of the two subsets of the group, though in other contexts the term *product* may stand for the set $A \times B$.

(10.1) **Lemma.** Let N be a normal subgroup of a group G. Then the product of two cosets aN, bN is again a coset, in fact

$$(aN)(bN) = abN.$$

Proof. Note that $Nb = bN$, by (6.18), and since N is a subgroup $NN = N$. The following formal manipulation proves the lemma:

$$(aN)(bN) = a(Nb)N = a(bN)N = abNN = abN. \square$$

This lemma allows us to define multiplication of two cosets C_1, C_2 by this rule: C_1C_2 is the product set. To compute the product coset, take any elements $a \in C_1$ and $b \in C_2$, so that $C_1 = aN$ and $C_2 = bN$. Then $C_1C_2 = abN$ is the coset containing ab. This is the way addition of congruence classes was defined in the last section.

For example, consider the cosets of the unit circle N in $G = \mathbb{C}^\times$. As we saw in Section 5, its cosets are the concentric circles

$$C_r = \{z \mid |z| = r\}.$$

Formula (10.1) amounts to the assertion that if $|\alpha| = r$ and $|\beta| = s$, then $|\alpha\beta| = rs$:

$$C_rC_s = C_{rs}.$$

The assumption that N is a *normal* subgroup of G is crucial to (10.1). If H is not a normal subgroup of G, then there will be left cosets C_1, C_2 of H in G whose products do not lie in a single left coset. For to say H is not normal means there are elements $h \in H$ and $a \in G$ so that $aha^{-1} \notin H$. Then the set

$$(aH)(a^{-1}H) \tag{10.2}$$

does not lie in any left coset. It contains $a1a^{-1}1 = 1$, which is an element of H. So if the set (10.2) is contained in a coset, that coset must be $H = 1H$. But it also contains $aha^{-1}1$, which is not in H. $\square$

It is customary to denote the set of cosets of a normal subgroup N of G by the symbol

$$G/N = \text{set of cosets of } N \text{ in } G. \tag{10.3}$$

This agrees with the notation $\mathbb{Z}/n\mathbb{Z}$ introduced in Section 9. Another notation we will frequently use for the set of cosets is the bar notation:

$$G/N = \overline{G} \quad \text{and} \quad aN = \overline{a},$$

so that $\overline{a}$ denotes the coset containing a. This is natural when we want to consider the map

$$\pi: G \longrightarrow \overline{G} = G/N \quad \text{sending} \quad a \rightsquigarrow \overline{a} = aN. \tag{10.4}$$

(10.5) **Theorem.** With the law of composition defined above, $\overline{G} = G/N$ is a group, and the map π (10.4) is a homomorphism whose kernel is N.

The order of G/N is the index $[G : N]$ of N in G.

(10.6) **Corollary.** Every normal subgroup of a group G is the kernel of a homomorphism. $\square$

This corollary allows us to apply everything that we know about homomorphisms to improve our understanding of normal subgroups.

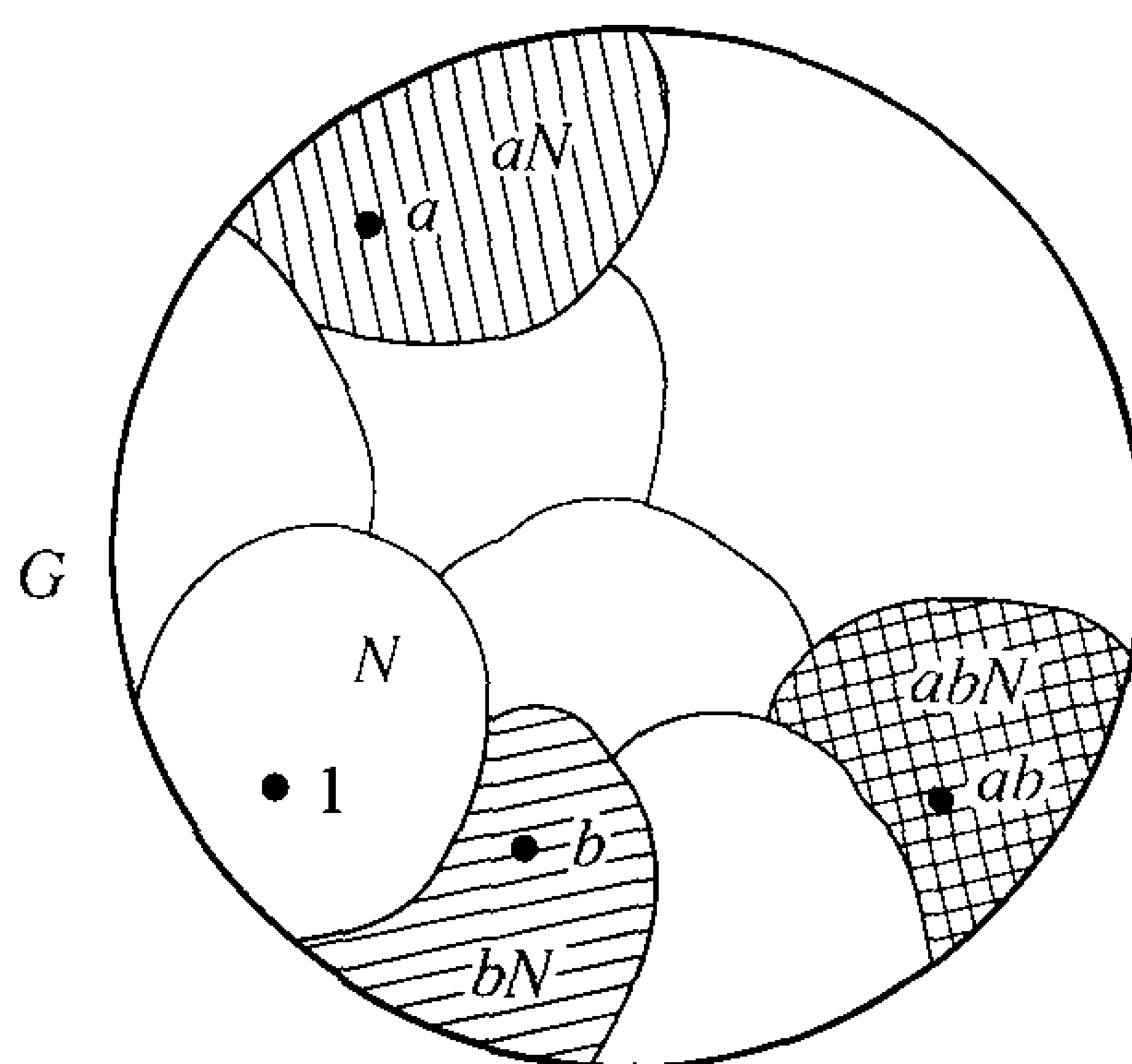
Proof of Theorem (10.5). First note that π is compatible with the laws of composition: Since multiplication of cosets is defined by multiplication of elements, $\pi(a)\pi(b) = \pi(ab)$. Moreover, the elements of G having the same image as the identity element 1 are those in N: $\overline{1} = 1N = N$. The group axioms in $\overline{G}$ follow from Lemma (10.7):

(10.7) **Lemma.** Let G be a group, and let S be any set with a law of composition. Let $\varphi: G \longrightarrow S$ be a surjective map which has the property $\varphi(a)\varphi(b) = \varphi(ab)$ for all a, b in G. Then S is a group.

Proof. Actually, any law concerning multiplication which holds in G will be carried over to S. The proof of the associative law is this: Let $s_1, s_2, s_3 \in S$. Since φ is surjective, we know that $s_i = \varphi(a_i)$ for some $a_i \in G$. Then

$$\begin{aligned}(s_1s_2)s_3 &= (\varphi(a_1)\varphi(a_2))\varphi(a_3) = \varphi(a_1a_2)\varphi(a_3) = \varphi(a_1a_2a_3)\\ &= \varphi(a_1)\varphi(a_2a_3) = \varphi(a_1)(\varphi(a_2)\varphi(a_3)) = s_1(s_2s_3).\end{aligned}$$

We leave the other group axioms as an exercise. □



(10.8) **Figure.** A schematic diagram of coset multiplication.

For example, let $G = \mathbb{R}^\times$ be the multiplicative group of nonzero real numbers, and let P be the subgroup of positive real numbers. There are two cosets, namely P and $-P$ = {negative reals}, and $\overline{G} = G/P$ is the group of two elements. The multiplication rule is the familiar rule: (*Neg*)(*Neg*) = (*Pos*), and so on.

The quotient group construction is related to a general homomorphism $\varphi: G \longrightarrow G'$ of groups as follows:

(10.9) **Theorem.** *First Isomorphism Theorem:* Let $\varphi: G \longrightarrow G'$ be a surjective group homomorphism, and let $N = \ker \varphi$. Then G/N is isomorphic to G' by the

map $\overline{\varphi}$ which sends the coset $\overline{a} = aN$ to $\varphi(a)$:

$$\overline{\varphi}(\overline{a}) = \varphi(a).$$

This is our fundamental method of identifying quotient groups. For example, the absolute value map $\mathbb{C}^\times \longrightarrow \mathbb{R}^\times$ maps the nonzero complex numbers to the positive real numbers, and its kernel is the unit circle U. So the quotient group $\mathbb{C}^\times/U$ is isomorphic to the multiplicative group of positive real numbers. Or, the determinant is a surjective homomorphism $GL_n(\mathbb{R}) \longrightarrow \mathbb{R}^\times$, whose kernel is the special linear group $SL_n(\mathbb{R})$. So the quotient $GL_n(\mathbb{R})/SL_n(\mathbb{R})$ is isomorphic to $\mathbb{R}^\times$.

Proof of the First Isomorphism Theorem. According to Proposition (5.13), the nonempty fibres of φ are the cosets aN. So we can think of $\overline{G}$ in either way, as the set of cosets or as the set of nonempty fibres of φ. Therefore the map we are looking for is the one defined in (5.10) for any map of sets. It maps $\overline{G}$ bijectively onto the image of φ, which is equal to G' because φ is surjective. By construction it is compatible with multiplication: $\overline{\varphi}(\overline{ab}) = \varphi(ab) = \varphi(a)\varphi(b) = \overline{\varphi}(\overline{a})\overline{\varphi}(\overline{b})$. □

Es giebt alſo ſehr viel verſchiedene Arten von Größen,
welche ſich nicht wohl herzehlen laßen;
und daher entſtehen die verſchiedene Theile der Mathematic,
deren eine jegliche mit einer beſondern Art von Größen beſchäftiget iſt.

Leonhard Euler

EXERCISES

1. The Definition of a Group

1. (a) Verify (1.17) and (1.18) by explicit computation.
   (b) Make a multiplication table for S_3.
2. (a) Prove that $GL_n(\mathbb{R})$ is a group.
   (b) Prove that S_n is a group.
3. Let S be a set with an associative law of composition and with an identity element. Prove that the subset of S consisting of invertible elements is a group.
4. Solve for y, given that $xyz^{-1}w = 1$ in a group.
5. Assume that the equation $xyz = 1$ holds in a group G. Does it follow that $yzx = 1$? That $yxz = 1$?
6. Write out all ways in which one can form a product of four elements a, b, c, d in the given order.
7. Let S be any set. Prove that the law of composition defined by $ab = a$ is associative.
8. Give an example of 2×2 matrices such that $A^{-1}B \neq BA^{-1}$.
9. Show that if $ab = a$ in a group, then $b = 1$, and if $ab = 1$, then $b = a^{-1}$.
10. Let a, b be elements of a group G. Show that the equation $ax = b$ has a unique solution in G.
11. Let G be a group, with multiplicative notation. We define an *opposite group* G^0 with law of composition $a \circ b$ as follows: The underlying set is the same as G, but the law of composition is the opposite; that is, we define $a \circ b = ba$. Prove that this defines a group.

2. Subgroups

1. Determine the elements of the cyclic group generated by the matrix $\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$ explicitly.
2. Let a, b be elements of a group G. Assume that a has order 5 and that $a^3b = ba^3$. Prove that $ab = ba$.
3. Which of the following are subgroups?
   (a) $GL_n(\mathbb{R}) \subset GL_n(\mathbb{C})$.
   (b) $\{1, -1\} \subset \mathbb{R}^\times$.
   (c) The set of positive integers in $\mathbb{Z}^+$.
   (d) The set of positive reals in $\mathbb{R}^\times$.
   (e) The set of all matrices $\begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}$, with $a \neq 0$, in $GL_2(\mathbb{R})$.
4. Prove that a nonempty subset H of a group G is a subgroup if for all $x, y \in H$ the element xy^{-1} is also in H.
5. An nth root of unity is a complex number z such that $z^n = 1$. Prove that the nth roots of unity form a cyclic subgroup of $\mathbb{C}^\times$ of order n.
6. (a) Find generators and relations analogous to (2.13) for the Klein four group.
   (b) Find all subgroups of the Klein four group.
7. Let a and b be integers.
   (a) Prove that the subset $a\mathbb{Z} + b\mathbb{Z}$ is a subgroup of $\mathbb{Z}^+$.
   (b) Prove that a and $b + 7a$ generate the subgroup $a\mathbb{Z} + b\mathbb{Z}$.
8. Make a multiplication table for the quaternion group H.
9. Let H be the subgroup generated by two elements a,b of a group G. Prove that if $ab = ba$, then H is an abelian group.
10. (a) Assume that an element x of a group has order rs. Find the order of x^r.
    (b) Assuming that x has arbitrary order n, what is the order of x^r?
11. Prove that in any group the orders of ab and of ba are equal.
12. Describe all groups G which contain no proper subgroup.
13. Prove that every subgroup of a cyclic group is cyclic.
14. Let G be a cyclic group of order n, and let r be an integer dividing n. Prove that G contains exactly one subgroup of order r.
15. (a) In the definition of subgroup, the identity element in H is required to be the identity of G. One might require only that H have an identity element, not that it is the same as the identity in G. Show that if H has an identity at all, then it is the identity in G, so this definition would be equivalent to the one given.
    (b) Show the analogous thing for inverses.
16. (a) Let G be a cyclic group of order 6. How many of its elements generate G?
    (b) Answer the same question for cyclic groups of order 5, 8, and 10.
    (c) How many elements of a cyclic group of order n are generators for that group?
17. Prove that a group in which every element except the identity has order 2 is abelian.
18. According to Chapter 1 (2.18), the elementary matrices generate $GL_n(\mathbb{R})$.
    (a) Prove that the elementary matrices of the first and third types suffice to generate this group.
    (b) The *special linear group* $SL_n(\mathbb{R})$ is the set of real $n \times n$ matrices whose determinant is 1. Show that $SL_n(\mathbb{R})$ is a subgroup of $GL_n(\mathbb{R})$.

*(c) Use row reduction to prove that the elementary matrices of the first type generate $SL_n(\mathbb{R})$. Do the 2×2 case first.

19. Determine the number of elements of order 2 in the symmetric group S_4.

20. (a) Let a, b be elements of an abelian group of orders m, n respectively. What can you say about the order of their product ab?

*(b) Show by example that the product of elements of finite order in a nonabelian group need not have finite order.

21. Prove that the set of elements of finite order in an abelian group is a subgroup.

22. Prove that the greatest common divisor of a and b, as defined in the text, can be obtained by factoring a and b into primes and collecting the common factors.

3. Isomorphisms

1. Prove that the additive group $\mathbb{R}^+$ of real numbers is isomorphic to the multiplicative group P of positive reals.

2. Prove that the products ab and ba are conjugate elements in a group.

3. Let a, b be elements of a group G, and let $a' = bab^{-1}$. Prove that $a = a'$ if and only if a and b commute.

4. (a) Let $b' = aba^{-1}$. Prove that $b'^n = ab^n a^{-1}$.

(b) Prove that if $aba^{-1} = b^2$, then $a^3ba^{-3} = b^8$.

5. Let $\varphi: G \longrightarrow G'$ be an isomorphism of groups. Prove that the inverse function φ^{-1} is also an isomorphism.

6. Let $\varphi: G \longrightarrow G'$ be an isomorphism of groups, let $x, y \in G$, and let $x' = \varphi(x)$ and $y' = \varphi(y)$.

(a) Prove that the orders of x and of x' are equal.

(b) Prove that if $xyx = yxy$, then $x'y'x' = y'x'y'$.

(c) Prove that $\varphi(x^{-1}) = x'^{-1}$.

7. Prove that the matrices $\begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}, \begin{bmatrix} 1 & \\ 1 & 1 \end{bmatrix}$ are conjugate elements in the group $GL_2(\mathbb{R})$ but that they are not conjugate when regarded as elements of $SL_2(\mathbb{R})$.

8. Prove that the matrices $\begin{bmatrix} 1 & \\ & 2 \end{bmatrix}, \begin{bmatrix} 1 & 3 \\ & 2 \end{bmatrix}$ are conjugate in $GL_2(\mathbb{R})$.

9. Find an isomorphism from a group G to its opposite group G^0 (Section 2, exercise 12).

10. Prove that the map $A \rightsquigarrow (A^t)^{-1}$ is an automorphism of $GL_n(\mathbb{R})$.

11. Prove that the set Aut G of automorphisms of a group G forms a group, the law of composition being composition of functions.

12. Let G be a group, and let $\varphi: G \longrightarrow G$ be the map $\varphi(x) = x^{-1}$.

(a) Prove that φ is bijective.

(b) Prove that φ is an automorphism if and only if G is abelian.

13. (a) Let G be a group of order 4. Prove that every element of G has order 1, 2, or 4.

(b) Classify groups of order 4 by considering the following two cases:

(i) G contains an element of order 4.

(ii) Every element of G has order < 4.

14. Determine the group of automorphisms of the following groups.

(a) $\mathbb{Z}^+$, (b) a cyclic group of order 10, (c) S_3.

15. Show that the functions $f = 1/x$, $g = (x - 1)/x$ generate a group of functions, the law of composition being composition of functions, which is isomorphic to the symmetric group S_3.

16. Give an example of two isomorphic groups such that there is more than one isomorphism between them.

4. Homomorphisms

1. Let G be a group, with law of composition written $x \# y$. Let H be a group with law of composition $u \circ v$. What is the condition for a map $\varphi: G \longrightarrow H'$ to be a homomorphism?

2. Let $\varphi: G \longrightarrow G'$ be a group homomorphism. Prove that for any elements $a_1, \ldots, a_k$ of G, $\varphi(a_1 \cdots a_k) = \varphi(a_1) \cdots \varphi(a_k)$.

3. Prove that the kernel and image of a homomorphism are subgroups.

4. Describe all homomorphisms $\varphi: \mathbb{Z}^+ \longrightarrow \mathbb{Z}^+$, and determine which are injective, which are surjective, and which are isomorphisms.

5. Let G be an abelian group. Prove that the nth power map $\varphi: G \longrightarrow G$ defined by $\varphi(x) = x^n$ is a homomorphism from G to itself.

6. Let $f: \mathbb{R}^+ \longrightarrow \mathbb{C}^\times$ be the map $f(x) = e^{ix}$. Prove that f is a homomorphism, and determine its kernel and image.

7. Prove that the absolute value map $| \ |: \mathbb{C}^\times \longrightarrow \mathbb{R}^\times$ sending $\alpha \rightsquigarrow |\alpha|$ is a homomorphism, and determine its kernel and image.

8. **(a)** Find all subgroups of S_3, and determine which are normal.
(b) Find all subgroups of the quaternion group, and determine which are normal.

9. **(a)** Prove that the composition $\varphi \circ \psi$ of two homomorphisms φ, ψ is a homomorphism.
(b) Describe the kernel of $\varphi \circ \psi$.

10. Let $\varphi: G \longrightarrow G'$ be a group homomorphism. Prove that $\varphi(x) = \varphi(y)$ if and only if $xy^{-1} \in \ker \varphi$.

11. Let G, H be cyclic groups, generated by elements x, y. Determine the condition on the orders m, n of x and y so that the map sending $x^i \rightsquigarrow y^i$ is a group homomorphism.

12. Prove that the $n \times n$ matrices M which have the block form $\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$ with $A \in GL_r(\mathbb{R})$ and $D \in GL_{n-r}(\mathbb{R})$ form a subgroup P of $GL_n(\mathbb{R})$, and that the map $P \longrightarrow GL_r(\mathbb{R})$ sending $M \rightsquigarrow A$ is a homomorphism. What is its kernel?

13. **(a)** Let H be a subgroup of G, and let $g \in G$. The *conjugate subgroup* gHg^{-1} is defined to be the set of all conjugates ghg^{-1}, where $h \in H$. Prove that gHg^{-1} is a subgroup of G.
(b) Prove that a subgroup H of a group G is normal if and only if $gHg^{-1} = H$ for all $g \in G$.

14. Let N be a normal subgroup of G, and let $g \in G$, $n \in N$. Prove that $g^{-1}ng \in N$.

15. Let φ and ψ be two homomorphisms from a group G to another group G', and let $H \subset G$ be the subset $\{x \in G \mid \varphi(x) = \psi(x)\}$. Prove or disprove: H is a subgroup of G.

16. Let $\varphi: G \longrightarrow G'$ be a group homomorphism, and let $x \in G$ be an element of order r. What can you say about the order of $\varphi(x)$?

17. Prove that the center of a group is a normal subgroup.

18. Prove that the center of $GL_n(\mathbb{R})$ is the subgroup $Z = \{cI \mid c \in \mathbb{R}, c \neq 0\}$.

19. Prove that if a group contains exactly one element of order 2, then that element is in the center of the group.

20. Consider the set U of real 3×3 matrices of the form

$$\begin{bmatrix} 1 & * & * \\ & 1 & * \\ & & 1 \end{bmatrix}.$$

(a) Prove that U is a subgroup of $SL_n(\mathbb{R})$.
(b) Prove or disprove: U is normal.
*(c) Determine the center of U.

21. Prove by giving an explicit example that $GL_2(\mathbb{R})$ is not a normal subgroup of $GL_2(\mathbb{C})$.

22. Let $\varphi: G \longrightarrow G'$ be a surjective homomorphism.
(a) Assume that G is cyclic. Prove that G' is cyclic.
(b) Assume that G is abelian. Prove that G' is abelian.

23. Let $\varphi: G \longrightarrow G'$ be a surjective homomorphism, and let N be a normal subgroup of G. Prove that $\varphi(N)$ is a normal subgroup of G'.

5. Equivalence Relations and Partitions

1. Prove that the nonempty fibres of a map form a partition of the domain.

2. Let S be a set of groups. Prove that the relation $G \sim H$ if G is isomorphic to H is an equivalence relation on S.

3. Determine the number of equivalence relations on a set of five elements.

4. Is the intersection $R \cap R'$ of two equivalence relations $R, R' \subset S \times S$ an equivalence relation? Is the union?

5. Let H be a subgroup of a group G. Prove that the relation defined by the rule $a \sim b$ if $b^{-1}a \in H$ is an equivalence relation on G.

6. (a) Prove that the relation x conjugate to y in a group G is an equivalence relation on G.
(b) Describe the elements a whose conjugacy class (= equivalence class) consists of the element a alone.

7. Let R be a relation on the set $\mathbb{R}$ of real numbers. We may view R as a subset of the (x, y)-plane. Explain the geometric meaning of the reflexive and symmetric properties.

8. With each of the following subsets R of the (x, y)-plane, determine which of the axioms (5.2) are satisfied and whether or not R is an equivalence relation on the set $\mathbb{R}$ of real numbers.
(a) $R = \{(s,s) \mid s \in \mathbb{R}\}$.
(b) $R =$ empty set.
(c) $R =$ locus $\{y = 0\}$.
(d) $R =$ locus $\{xy + 1 = 0\}$.
(e) $R =$ locus $\{x^2y - xy^2 - x + y = 0\}$.
(f) $R =$ locus $\{x^2 - xy + 2x - 2y = 0\}$.

9. Describe the smallest equivalence relation on the set of real numbers which contains the line $x - y = 1$ in the (x, y)-plane, and sketch it.

10. Draw the fibres of the map from the (x,z)-plane to the y-axis defined by the map $y = zx$.

11. Work out rules, obtained from the rules on the integers, for addition and multiplication on the set (5.8).

12. Prove that the cosets (5.14) are the fibres of the map φ.

6. Cosets

1. Determine the index $[\mathbb{Z} : n\mathbb{Z}]$.

2. Prove directly that distinct cosets do not overlap.

3. Prove that every group whose order is a power of a prime p contains an element of order p.

4. Give an example showing that left cosets and right cosets of $GL_2(\mathbb{R})$ in $GL_2(\mathbb{C})$ are not always equal.

5. Let H, K be subgroups of a group G of orders 3, 5 respectively. Prove that $H \cap K = \{1\}$.

6. *Justify (6.15) carefully.*

7. **(a)** Let G be an abelian group of odd order. Prove that the map $\varphi: G \longrightarrow G$ defined by $\varphi(x) = x^2$ is an automorphism.

(b) Generalize the result of (a).

8. Let W be the additive subgroup of $\mathbb{R}^m$ of solutions of a system of homogeneous linear equations $AX = 0$. Show that the solutions of an inhomogeneous system $AX = B$ form a coset of W.

9. Let H be a subgroup of a group G. Prove that the number of left cosets is equal to the number of right cosets **(a)** if G is finite and **(b)** in general.

10. **(a)** Prove that every subgroup of index 2 is normal.

(b) Give an example of a subgroup of index 3 which is not normal.

11. Classify groups of order 6 by analyzing the following three cases.

(a) G contains an element of order 6.

(b) G contains an element of order 3 but none of order 6.

(c) All elements of G have order 1 or 2.

12. Let G, H be the following subgroups of $GL_2(\mathbb{R})$:

$$G = \left\{\begin{bmatrix} x & y \\ 0 & 1 \end{bmatrix}\right\}, H = \left\{\begin{bmatrix} x & 0 \\ 0 & 1 \end{bmatrix}\right\}, x > 0.$$

An element of G can be represented by a point in the (x, y)-plane. Draw the partitions of the plane into left and into right cosets of H.

7. Restriction of a Homomorphism to a Subgroup

1. Let G and G' be finite groups whose orders have no common factor. Prove that the only homomorphism $\varphi: G \longrightarrow G'$ is the trivial one $\varphi(x) = 1$ for all x.

2. Give an example of a permutation of even order which is odd and an example of one which is even.

3. **(a)** Let H and K be subgroups of a group G. Prove that the intersection $xH \cap yK$ of two cosets of H and K is either empty or else is a coset of the subgroup $H \cap K$.

(b) Prove that if H and K have finite index in G then $H \cap K$ also has finite index.

4. Prove Proposition (7.1).

5. Let H, N be subgroups of a group G, with N normal. Prove that $HN = NH$ and that this set is a subgroup.

6. Let $\varphi: G \longrightarrow G'$ be a group homomorphism with kernel K, and let H be another subgroup of G. Describe $\varphi^{-1}(\varphi(H))$ in terms of H and K.

7. Prove that a group of order 30 can have at most 7 subgroups of order 5.

***8.** Prove the *Correspondence Theorem:* Let $\varphi: G \longrightarrow G'$ be a surjective group homomorphism with kernel N. The set of subgroups H' of G' is in bijective correspondence with the set of subgroups H of G which contain N, the correspondence being defined by the maps $H \rightsquigarrow \varphi(H)$ and $\varphi^{-1}(H') \leftarrow\!\!\!\sim H'$. Moreover, normal subgroups of G correspond to normal subgroups of G'.

9. Let G and G' be cyclic groups of orders 12 and 6 generated by elements x, y respectively, and let $\varphi: G \longrightarrow G'$ be the map defined by $\varphi(x^i) = y^i$. Exhibit the correspondence referred to the previous problem explicitly.

8. Products of Groups

1. Let G, G' be groups. What is the order of the product group $G \times G'$?

2. Is the symmetric group S_3 a direct product of nontrivial groups?

3. Prove that a finite cyclic group of order rs is isomorphic to the product of cyclic groups of orders r and s if and only if r and s have no common factor.

4. In each of the following cases, determine whether or not G is isomorphic to the product of H and K.

(a) $G = \mathbb{R}^\times$, $H = \{\pm 1\}$, $K = \{\text{positive real numbers}\}$.

(b) $G = \{\text{invertible upper triangular } 2 \times 2 \text{ matrices}\}$, $H = \{\text{invertible diagonal matrices}\}$, $K = \{\text{upper triangular matrices with diagonal entries } 1\}$.

(c) $G = \mathbb{C}^\times$ and $H = \{\text{unit circle}\}$, $K = \{\text{positive reals}\}$.

5. Prove that the product of two infinite cyclic groups is not infinite cyclic.

6. Prove that the center of the product of two groups is the product of their centers.

7. **(a)** Let H, K be subgroups of a group G. Show that the set of products $HK = \{hk \mid h \in H, k \in K\}$ is a subgroup if and only if $HK = KH$.

(b) Give an example of a group G and two subgroups H, K such that HK is not a subgroup.

8. Let G be a group containing normal subgroups of orders 3 and 5 respectively. Prove that G contains an element of order 15.

9. Let G be a finite group whose order is a product of two integers: $n = ab$. Let H, K be subgroups of G of orders a and b respectively. Assume that $H \cap K = \{1\}$. Prove that $HK = G$. Is G isomorphic to the product group $H \times K$?

10. Let $x \in G$ have order m, and let $y \in G'$ have order n. What is the order of (x, y) in $G \times G'$?

11. Let H be a subgroup of a group G, and let $\varphi: G \longrightarrow H$ be a homomorphism whose restriction to H is the identity map: $\varphi(h) = h$, if $h \in H$. Let $N = \ker \varphi$.

(a) Prove that if G is abelian then it is isomorphic to the product group $H \times N$.

(b) Find a bijective map $G \longrightarrow H \times N$ without the assumption that G is abelian, but show by an example that G need not be isomorphic to the product group.

9. Modular Arithmetic

1. Compute $(7 + 14)(3 - 16)$ modulo 17.
2. (a) Prove that the square a^2 of an integer a is congruent to 0 or 1 modulo 4.
   (b) What are the possible values of a^2 modulo 8?
3. (a) Prove that 2 has no inverse modulo 6.
   (b) Determine all integers n such that 2 has an inverse modulo n.
4. Prove that every integer a is congruent to the sum of its decimal digits modulo 9.
5. Solve the congruence $2x \equiv 5$ (a) modulo 9 and (b) modulo 6.
6. Determine the integers n for which the congruences $x + y \equiv 2$, $2x - 3y \equiv 3$ (modulo n) have a solution.
7. Prove the associative and commutative laws for multiplication in $\mathbb{Z}/n\mathbb{Z}$.
8. Use Proposition (2.6) to prove the *Chinese Remainder Theorem:* Let m, n, a, b be integers, and assume that the greatest common divisor of m and n is 1. Then there is an integer x such that $x \equiv a$ (modulo m) and $x \equiv b$ (modulo n).

10. Quotient Groups

1. Let G be the group of invertible real upper triangular 2×2 matrices. Determine whether or not the following conditions describe normal subgroups H of G. If they do, use the First Isomorphism Theorem to identify the quotient group G/H.
   (a) $a_{11} = 1$ (b) $a_{12} = 0$ (c) $a_{11} = a_{22}$ (d) $a_{11} = a_{22} = 1$
2. Write out the proof of (10.1) in terms of elements.
3. Let P be a partition of a group G with the property that for any pair of elements A, B of the partition, the product set AB is contained entirely within another element C of the partition. Let N be the element of P which contains 1. Prove that N is a normal subgroup of G and that P is the set of its cosets.
4. (a) Consider the presentation (1.17) of the symmetric group S_3. Let H be the subgroup $\{1, y\}$. Compute the product sets $(1H)(xH)$ and $(1H)(x^2H)$, and verify that they are not cosets.
   (b) Show that a cyclic group of order 6 has two generators satisfying the rules $x^3 = 1$, $y^2 = 1$, $yx = xy$.
   (c) Repeat the computation of (a), replacing the relations (1.18) by the relations given in part (b). Explain.
5. Identify the quotient group $\mathbb{R}^\times/P$, where P denotes the subgroup of positive real numbers.
6. Let $H = \{\pm 1, \pm i\}$ be the subgroup of $G = \mathbb{C}^\times$ of fourth roots of unity. Describe the cosets of H in G explicitly, and prove that G/H is isomorphic to G.
7. Find all normal subgroups N of the quaternion group H, and identify the quotients H/N.
8. Prove that the subset H of $G = GL_n(\mathbb{R})$ of matrices whose determinant is positive forms a normal subgroup, and describe the quotient group G/H.
9. Prove that the subset $G \times 1$ of the product group $G \times G'$ is a normal subgroup isomorphic to G and that $(G \times G')/(G \times 1)$ is isomorphic to G'.
10. Describe the quotient groups $\mathbb{C}^\times/P$ and $\mathbb{C}^\times/U$, where U is the subgroup of complex numbers of absolute value 1 and P denotes the positive reals.
11. Prove that the groups $\mathbb{R}^+/\mathbb{Z}^+$ and $\mathbb{R}^+/2\pi\mathbb{Z}^+$ are isomorphic.

Miscellaneous Problems

1. What is the product of all mth roots of unity in $\mathbb{C}$?

2. Compute the group of automorphisms of the quaternion group.

3. Prove that a group of even order contains an element of order 2.

4. Let $K \subset H \subset G$ be subgroups of a finite group G. Prove the formula $[G : K] = [G : H][H : K]$.

***5.** A *semigroup* S is a set with an associative law of composition and with an identity. But elements are not required to have inverses, so the cancellation law need not hold. The semigroup S is said to be generated by an element s if the set $\{1, s, s^2, \ldots\}$ of nonnegative powers of s is the whole set S. For example, the relations $s^2 = 1$ and $s^2 = s$ describe two different semigroup structures on the set $\{1, s\}$. Define isomorphism of semigroups, and describe all isomorphism classes of semigroups having a generator.

6. Let S be a semigroup with finitely many elements which satisfies the Cancellation Law (1.12). Prove that S is a group.

***7.** Let $a = (a_1, \ldots, a_k)$ and $b = (b_1, \ldots, b_k)$ be points in k-dimensional space $\mathbb{R}^k$. A *path* from a to b is a continuous function on the interval $[0, 1]$ with values in $\mathbb{R}^k$, that is, a function $f\colon [0, 1] \longrightarrow \mathbb{R}^k$, sending $t \rightsquigarrow f(t) = (x_1(t), \ldots, x_k(t))$, such that $f(0) = a$ and $f(1) = b$. If S is a subset of $\mathbb{R}^k$ and if $a, b \in S$, we define $a \sim b$ if a and b can be joined by a path lying entirely in S.

(a) Show that this is an equivalence relation on S. Be careful to check that the paths you construct stay within the set S.

(b) A subset S of $\mathbb{R}^k$ is called *path connected* if $a \sim b$ for any two points $a, b \in S$. Show that every subset S is partitioned into path-connected subsets with the property that two points in different subsets can not be connected by a path in S.

(c) Which of the following loci in $\mathbb{R}^2$ are path-connected? $\{x^2 + y^2 = 1\}$, $\{xy = 0\}$, $\{xy = 1\}$.

***8.** The set of $n \times n$ matrices can be identified with the space $\mathbb{R}^{n \times n}$. Let G be a subgroup of $GL_n(\mathbb{R})$. Prove each of the following.

(a) If $A, B, C, D \in G$, and if there are paths in G from A to B and from C to D, then there is a path in G from AC to BD.

(b) The set of matrices which can be joined to the identity I forms a normal subgroup of G (called the *connected component* of G).

***9.** (a) Using the fact that $SL_n(\mathbb{R})$ is generated by elementary matrices of the first type (see exercise 18, Section 2), prove that this group is path-connected.

(b) Show that $GL_n(\mathbb{R})$ is a union of two path-connected subsets, and describe them.

10. Let H, K be subgroups of a group G, and let $g \in G$. The set

$$HgK = \{x \in G \mid x = hgk \text{ for some } h \in H, k \in K\}$$

is called a *double coset*.

(a) Prove that the double cosets partition G.

(b) Do all double cosets have the same order?

11. Let H be a subgroup of a group G. Show that the double cosets HgH are the left cosets gH if H is normal, but that if H is not normal then there is a double coset which properly contains a left coset.

***12.** Prove that the double cosets in $GL_n(\mathbb{R})$ of the subgroups $H = \{\text{lower triangular matrices}\}$ and $K = \{\text{upper triangular matrices}\}$ are the sets HPK, where P is a permutation matrix.

Chapter 3

Vector Spaces

Immer mit den einfachsten Beispielen anfangen.

David Hilbert

1. REAL VECTOR SPACES

The basic models for vector spaces are the spaces of n-dimensional row or column vectors:

$$\mathbb{R}^n\text{: the set of row vectors } v = (a_1, \ldots, a_n)\text{, or}$$

$$\text{the set of column vectors } v = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

Though row vectors take less space to write, the definition of matrix multiplication makes column vectors more convenient for us. So we will work with column vectors most of the time. To save space, we will occasionally write a column vector in the form $(a_1, \ldots, a_n)^t$.

For the present we will study only two operations:

$$(1.1) \qquad \textit{vector addition: } \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1+b_1 \\ \vdots \\ a_n+b_n \end{bmatrix}, \text{ and}$$

$$\textit{scalar multiplication: } c\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} ca_1 \\ \vdots \\ ca_n \end{bmatrix}.$$

These operations make $\mathbb{R}^n$ into a *vector space*. Before going to the formal definition of a vector space, let us look at some other examples—nonempty subsets of $\mathbb{R}^n$ closed under the operations (1.1). Such a subset is called a *subspace*.

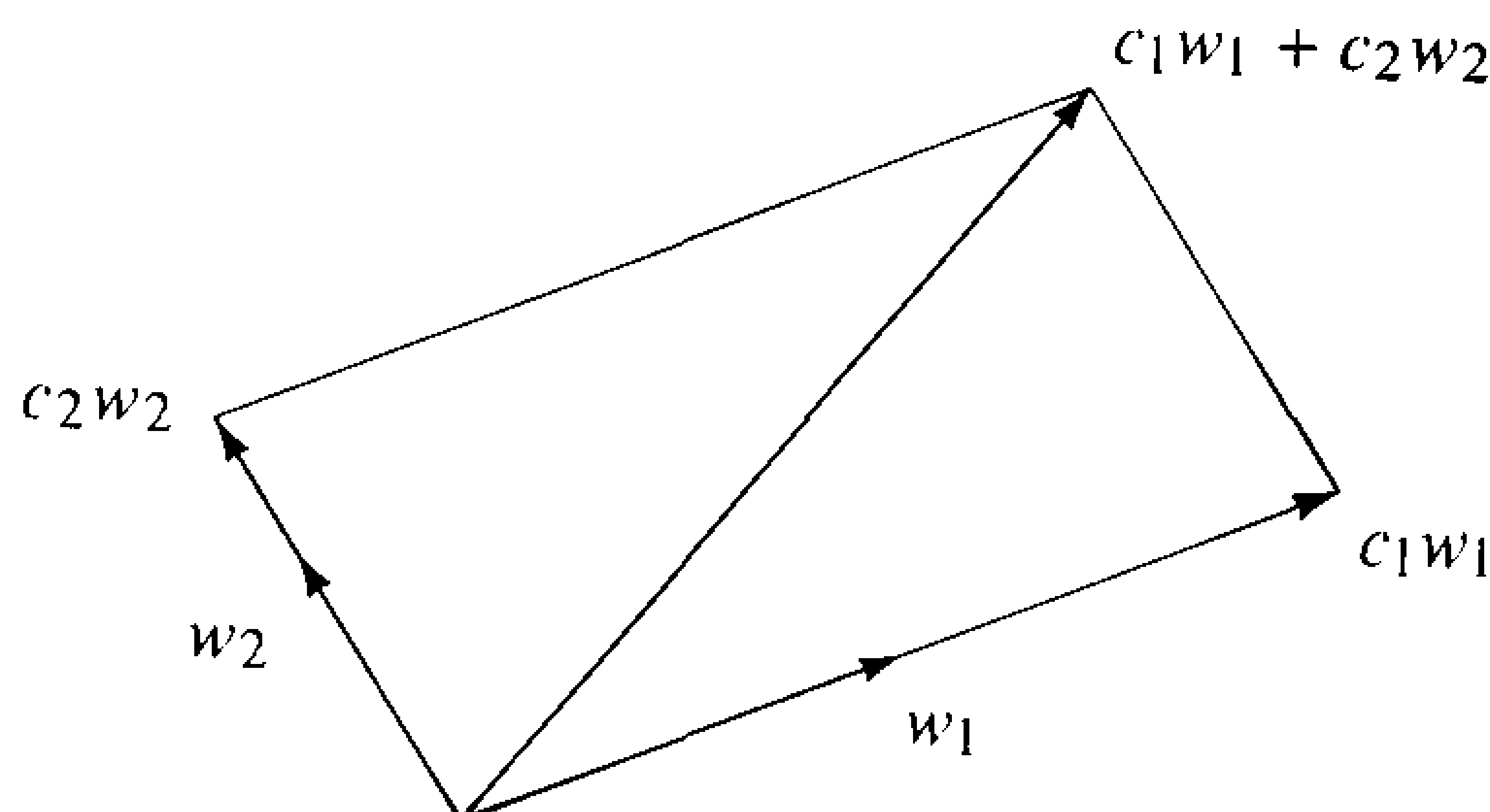
(1.2) **Example.** The subspaces W of the space $\mathbb{R}^2$ are of three types:

(i) the zero vector alone: $W = \{0\}$;
(ii) the vectors lying on a line L through the origin;
(iii) the whole space: $W = \mathbb{R}^2$.

This can be seen from the parallelogram law for addition of vectors. If W contains two vectors w_1, w_2 not lying on one line, then every vector v can be obtained from these two vectors as a "linear combination"

$$c_1w_1 + c_2w_2,$$

where c_1, c_2 are scalars. So $W = \mathbb{R}^2$ in this case. If W does not contain two such vectors, then we are in one of the remaining cases. □



Similarly, it can be shown that the subspaces of $\mathbb{R}^3$ are of four types:

(i) the zero vector;
(ii) the vectors lying on a line through the origin;
(iii) the vectors lying in a plane through the origin;
(iv) the whole space $\mathbb{R}^3$.

This classification of subspaces of $\mathbb{R}^2$ and $\mathbb{R}^3$ will be clarified in Section 4 by the concept of *dimension*.

Systems of homogeneous linear equations furnish many examples. The set of solutions of such a system is always a subspace. For, if we write the system in matrix notation as $AX = 0$, where A is an $m \times n$ matrix and X is a column vector, then it is clear that

(a) $AX = 0$ and $AY = 0$ imply $A(X + Y) = 0$. In other words, if X and Y are solutions, so is $X + Y$.
(b) $AX = 0$ implies $AcX = 0$: If X is a solution, so is cX.

For example, let W be the set of solutions of the equation

$$2x_1 - x_2 - 2x_3 = 0, \text{ or } AX = 0, \tag{1.3}$$

where $A = [2\ -1\ \ 2]$. This space is the set of vectors lying in the plane through the origin and orthogonal to A. Every solution is a linear combination $c_1w_1 + c_2w_2$ of two particular solutions w_1, w_2. Most pairs of solutions, for example

$$w_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad w_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \tag{1.4}$$

will span the space of solutions in this way. Thus every solution has the form

$$c_1w_1 + c_2w_2 = \begin{bmatrix} c_1+c_2 \\ 2c_2 \\ c_1 \end{bmatrix}, \tag{1.5}$$

where c_1, c_2 are arbitrary constants. Another choice of the particular solutions w_1, w_2 would result in a different but equivalent description of the space of all solutions.

(1.6) **Definition.** A *real vector space* is a set V together with two laws of composition:

(a) *Addition:* $V \times V \longrightarrow V$, written $v, w \rightsquigarrow v + w$

(b) *Scalar multiplication:* $\mathbb{R} \times V \longrightarrow V$, written $c, v \rightsquigarrow cv$

These laws of composition must satisfy the following axioms:

(i) Addition makes V into an abelian group V^+.

(ii) Scalar multiplication is associative with multiplication of real numbers:

$$(ab)v = a(bv).$$

(iii) Scalar multiplication by the real number 1 is the identity operation:

$$1v = v.$$

(iv) Two distributive laws hold:

$$(a + b)v = av + bv$$
$$a(v + w) = av + aw.$$

Of course all the axioms should be quantified universally; that is, they are assumed to hold for all $a, b \in \mathbb{R}$ and all $v, w \in V$.

The identity element for the addition law in V is denoted by 0, or by 0_V if there is danger of confusing the zero vector with the number zero.

Notice that scalar multiplication associates to every pair consisting of a real number c and a vector v another vector cv. Such a rule is called an *external law of composition* on the vector space.

Multiplication of two vectors is not a part of the structure, though various products, such as the cross product of vectors in $\mathbb{R}^3$, can be defined. These products aren't completely intrinsic; they depend on choosing coordinates. So they are considered to be additional structure on the vector space.

Read axiom (ii) carefully. The left side means multiply a and b as real numbers, then scalar multiply ab and v, to get a vector. On the right side, both operations are scalar multiplication.

The two laws of composition are related by the essential distributive laws. Note that in the first distributive law the symbol $+$ on the left stands for addition of real numbers, while on the right, it stands for addition of vectors.

(1.7) **Proposition.** The following identities hold in a vector space V:

(a) $0_{\mathbb{R}}v = 0_V$, for all $v \in V$,
(b) $c0_V = 0_V$, for all $c \in \mathbb{R}$,
(c) $(-1)v = -v$, for all $v \in V$.

Proof. To see (a), we use the distributive law to write

$$0v + 0v = (0 + 0)v = 0v = 0v + 0.$$

Cancelling $0v$ from both sides, we obtain $0v = 0$. Please go through this carefully, noting which symbols 0 refer to the number and which refer to the vector.

Similarly, $c0 + c0 = c(0 + 0) = c0$. Hence $c0 = 0$. Finally,

$$v + -1v = 1v + -1v = (1 + -1)v = 0v = 0.$$

Hence $-1v$ is the additive inverse of v. □

(1.8) **Examples.**

(a) A subspace of $\mathbb{R}^n$ is a vector space, with the laws of composition induced from those on $\mathbb{R}^n$.
(b) Let $V = \mathbb{C}$ be the set of complex numbers. Forget multiplication of complex numbers, and keep only addition $\alpha + \beta$ and multiplication $c\alpha$ of a complex number α by a real number c. These operations make $\mathbb{C}$ into a real vector space.
(c) The set of real polynomials $p(x) = a_n x^n + \cdots + a_0$ is a vector space, with addition of polynomials and multiplication of polynomials by scalars as its laws of composition.
(d) Let V be the set of continuous real-valued functions on the interval $[0, 1]$. Look only at the operations of addition of functions $f + g$ and multiplication of functions by numbers cf. This makes V a real vector space.

Note that each of our examples has more structure than we look at when we view it as a vector space. This is typical. Any particular example is sure to have some extra features which distinguish it from others, but this is not a drawback of the definition. On the contrary, the strength of the abstract approach lies in the fact that consequences of the general axioms can be applied to many different examples.

2. ABSTRACT FIELDS

It is convenient to treat the real and complex cases simultaneously in linear algebra. This can be done by listing the properties of the "scalars" which are needed axiomatically, and doing so leads to the notion of a *field*.

It used to be customary to speak only of subfields of the complex numbers. A *subfield* of $\mathbb{C}$ is any subset which is closed under the four operations addition, subtraction, multiplication, and division, and which contains 1. In other words, F is a subfield of $\mathbb{C}$ if the following properties hold:

(2.1)

(a) If $a, b \in F$, then $a + b \in F$.
(b) If $a \in F$, then $-a \in F$.
(c) If $a, b \in F$, then $ab \in F$.
(d) If $a \in F$ and $a \neq 0$, then $a^{-1} \in F$.
(e) $1 \in F$.

Note that we can use axioms (a), (b), and (e) to conclude that $1 - 1 = 0$ is an element of F. Thus F is a subset which is a subgroup of $\mathbb{C}^+$ under addition and such that $F - \{0\} = F^\times$ is a subgroup of $\mathbb{C}^\times$ under multiplication. Conversely, any such subset is a subfield.

Here are some examples of subfields of $\mathbb{C}$:

(2.2) **Examples.**

(a) $F = \mathbb{R}$, the field of real numbers.
(b) $F = \mathbb{Q}$, the field of rational numbers (= fractions of integers).
(c) $F = \mathbb{Q}[\sqrt{2}]$, the field of all complex numbers of the form $a + b\sqrt{2}$, where $a, b \in \mathbb{Q}$.

It is a good exercise to check axioms (2.1) for the last example.

These days, it is customary to introduce fields abstractly. The notion of an abstract field is harder to grasp than that of a subfield of $\mathbb{C}$, but it contains important new classes of fields, including finite fields.

(2.3) **Definition.** A *field* F is a set together with two laws of composition

$$F \times F \xrightarrow{+} F \quad \text{and} \quad F \times F \xrightarrow{\times} F$$
$$a, b \rightsquigarrow a + b \qquad a, b \rightsquigarrow ab$$

called addition and multiplication, and satisfying the following axioms:

(i) Addition makes F into an abelian group F^+. Its identity element is denoted by 0.
(ii) Multiplication is associative and commutative and makes $F^\times = F - \{0\}$ into a group. Its identity element is denoted by 1.
(iii) Distributive law: For all $a, b, c \in F$, $(a + b)c = ac + bc$.

The first two axioms describe properties of the two laws of composition, addition and multiplication, separately. The third axiom, the distributive law, is the one which relates addition to multiplication. This axiom is crucial, because if the two laws were unrelated, we could just as well study each of them separately. Of course we know that the real numbers satisfy these axioms, but the fact that they are all that is needed for arithmetic operations can only be understood after some experience in working with them.

One can operate with matrices A whose entries a_{ij} are in any field F. The discussion of Chapter 1 can be repeated without change, and you should go back to look at this material again with this in mind.

The simplest examples of fields besides the subfields of the complex numbers are certain finite fields called the prime fields, which we will now describe. We saw in Section 9 of Chapter 2 that the set $\mathbb{Z}/n\mathbb{Z}$ of congruence classes modulo n has laws of addition and multiplication derived from addition and multiplication of integers. Now all of the axioms for a field hold for the integers, except for the existence of multiplicative inverses in axiom (2.3ii). The integers are not closed under division. And as we have already remarked, such axioms carry over to addition and multiplication of congruence classes. But there is no reason to suppose that multiplicative inverses will exist for congruence classes, and in fact they need not. The class of 2, for example, does not have a multiplicative inverse modulo 6. So it is a surprising fact that if p is a prime integer then all nonzero congruence classes modulo p have inverses, and therefore the set $\mathbb{Z}/p\mathbb{Z}$ is a field. This field is called a *prime field* and is usually denoted by $\mathbb{F}_p$:

(2.4)
$$\mathbb{F}_p = \{\overline{0}, \overline{1}, \ldots, \overline{p-1}\} = \mathbb{Z}/p\mathbb{Z}.$$

(2.5) **Theorem.** Let p be a prime integer. Every nonzero congruence class $\overline{a}$ (modulo p) has a multiplicative inverse, and hence $\mathbb{F}_p$ is a field with p elements.

The theorem can also be stated as follows:

(2.6) *Let p be a prime, and let a be any integer not divisible by p. There is an integer b such that $ab \equiv 1$ (modulo p).*

For $ab \equiv 1$ (modulo p) is the same as $\overline{a}\overline{b} = \overline{ab} = \overline{1}$, which means that $\overline{b}$ is the multiplicative inverse of $\overline{a}$.

For example, let $p = 13$ and $\overline{a} = \overline{6}$. Then $\overline{a}^{-1} = \overline{11}$ because

$$6 \cdot 11 = 66 \equiv 1 \text{ (modulo 13).}$$

Finding the inverse of a congruence class $\overline{a}$ (modulo p) is not easy in general, but it can be done by trial and error if p is small. A systematic way is to compute the powers of $\overline{a}$. Since every nonzero congruence class has an inverse, the set of all of them forms a finite group of order $p - 1$, usually denoted by $\mathbb{F}_p^{\times}$. So every element $\overline{a}$ has finite order dividing $p - 1$. Thus if $p = 13$ and $\overline{a} = \overline{3}$, we find $\overline{a}^2 = \overline{9}$, and $\overline{a}^3 = \overline{27} = \overline{1}$, which shows that $\overline{a}$ has order 3. We are lucky: $\overline{a}^{-1} = \overline{a}^2 = \overline{9}$. On the other hand, if we had tried this method with $\overline{a} = \overline{6}$, we would have found that $\overline{6}$ has order 12. The computation would have been lengthy.

Proof of Theorem (2.5). Let $\overline{a} \in \mathbb{F}_p$ be any nonzero element, and let us use the method just discussed to show that $\overline{a}$ has an inverse. We consider the powers $1, \overline{a}, \overline{a}^2, \overline{a}^3, \ldots$. Since there are infinitely many powers and only finitely many elements in $\mathbb{F}_p$, there must be two powers which are equal, say $\overline{a}^m = \overline{a}^n$, where $m < n$. At this point, we would like to cancel $\overline{a}^m$, to obtain $\overline{1} = \overline{a}^{n-m}$. Once this cancellation is justified, we will have shown that $\overline{a}^{n-m-1}$ is the inverse of $\overline{a}$. This will complete the proof.

Here is the cancellation law we need:

(2.7) **Lemma.** *Cancellation Law:* Let $\overline{a}, \overline{c}, \overline{d}$ be elements of $\mathbb{F}_p$ with $\overline{a} \neq \overline{0}$. If $\overline{a}\overline{c} = \overline{a}\overline{d}$, then $\overline{c} = \overline{d}$.

Proof. Set $\overline{b} = \overline{c} - \overline{d}$. Then the statement of the lemma becomes: If $\overline{a}\overline{b} = \overline{0}$ and $\overline{a} \neq 0$, then $\overline{b} = 0$. To prove this, we represent the congruence classes $\overline{a}, \overline{b}$ by integers a, b. Then what has to be shown is the following intuitively plausible fact:

(2.8) **Lemma.** Let p be a prime integer and let a, b be integers. If p divides the product ab, then p divides a or p divides b.

Proof. Suppose that p does not divide a, but that p divides ab. We must show that p divides b. Since p is a prime, 1 and p are the only positive integers which divide it. Since p does not divide a, the only common divisor of p and a is 1. So 1 is their greatest common divisor. By Proposition (2.6) of Chapter 2, there are integers r, s so that $1 = rp + sa$. Multiply both sides by b: $b = rpb + sab$. Both of the terms on the right side of this equality are divisible by p; hence the left side a is divisible by p too, as was to be shown. □

As with congruences in general, computations in the field $\mathbb{F}_p$ can be made by working with integers, except that division can not be carried out in the integers. This difficulty can often be handled by putting everything on a common denominator in such a way that the required division is left until the end. For example, suppose we ask for solutions of a system of n linear equations in n unknowns, in the field $\mathbb{F}_p$.

We represent the system of equations by an integer system, choosing representatives for the residue classes in a convenient way. Say that the integer system is $AX = B$, where A is an $n \times n$ integer matrix and B is an integer column vector. Then to solve the system in $\mathbb{F}_p$, we try to invert the matrix A modulo p. Cramer's Rule, $(\text{adj } A)A = \delta I$, where $\delta = \det A$, is a formula valid in the integers [Chapter 1 (5.7)], and therefore it also holds in $\mathbb{F}_p$ when the matrix entries are replaced by their congruence classes. If the residue class of δ is not zero, then we can invert the matrix A in $\mathbb{F}_p$ by computing $\delta^{-1}(\text{adj } A)$.

(2.9) **Corollary.** Consider a system $AX = B$ of n linear equations in n unknowns where the entries of A, B are in $\mathbb{F}_p$. The system has a unique solution in $\mathbb{F}_p$ if $\det A \neq 0$ in $\mathbb{F}_p$. □

For example, consider the system of linear equations $AX = B$, where

$$A = \begin{bmatrix} 8 & 3 \\ 2 & 6 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 3 \\ -1 \end{bmatrix}.$$

Since the coefficients are integers, they define a system of equations in $\mathbb{F}_p$ for any prime p. The determinant of A is 42, so the system has a unique solution in $\mathbb{F}_p$ for all p different from 2, 3 and 7. Thus if $p = 13$, we find $\det A = 3$ when evaluated (modulo 13). We already saw that $3^{-1} = 9$ in $\mathbb{F}_{13}$. So we can use Cramer's Rule to compute

$$A^{-1} = \begin{bmatrix} 2 & -1 \\ 8 & 7 \end{bmatrix} \quad \text{and} \quad X = A^{-1}B = \begin{bmatrix} 7 \\ 4 \end{bmatrix}, \text{ in } \mathbb{F}_{13}.$$

The system has no solution in $\mathbb{F}_2$ or $\mathbb{F}_3$. It happens to have solutions in $\mathbb{F}_7$, though $\det A = 0$ in that field.

We remark in passing that invertible matrices with entries in the field $\mathbb{F}_p$ provide new examples of finite groups—the general linear groups over finite fields:

$$GL_n(\mathbb{F}_p) = \{n \times n \text{ invertible matrices with entries in } \mathbb{F}_p\}.$$

The smallest of these is the group $GL_2(\mathbb{F}_2)$ of invertible 2×2 matrices with entries (modulo 2), which consists of the six matrices

(2.10)

$$GL_2(\mathbb{F}_2) = \left\{ \begin{bmatrix} 1 & \\ & 1 \end{bmatrix}, \begin{bmatrix} & 1 \\ 1 & \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & \end{bmatrix}, \begin{bmatrix} 1 & \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} & 1 \\ 1 & 1 \end{bmatrix} \right\}.$$

There is one property of the finite fields $F = \mathbb{F}_p$ which distinguishes them from subfields of $\mathbb{C}$ and which affects computations occasionally. This property is that adding 1 to itself a certain number of times (in fact p times) gives 0. A field F is said to have *characteristic* p if $1 + \cdots + 1$ (p terms) $= 0$ in F, and if p is the smallest positive integer with that property. In other words, the characteristic of F is the order of 1, as an element of the additive group F^+, provided that the order is finite (Chapter 2, Section 2). In case the order is infinite, that is, $1 + \cdots + 1$ is

never 0 in F, the field is, paradoxically, said to have *characteristic zero*. Thus subfields of $\mathbb{C}$ have characteristic zero, while the prime field $\mathbb{F}_p$ has characteristic p. It can be shown that the characteristic of any field F is either zero or a prime number.

Now let F be an arbitrary field. A vector space over a field F is defined as in (1.6), with F replacing $\mathbb{R}$.

(2.11) **Definition.** A *vector space* V over a field F is a set together with two laws of composition:

(a) *addition:* $V \times V \longrightarrow V$, written $v, w \rightsquigarrow v + w$,
(b) *scalar multiplication:* $F \times V \longrightarrow V$, written $c, v \rightsquigarrow cv$,

and satisfying the following axioms:

(i) Addition makes V into a commutative group V^+.
(ii) Scalar multiplication is associative with multiplication in F:

$$(ab)v = a(bv), \text{ for all } a, b \in F \text{ and } v \in V.$$

(iii) The element 1 acts as identity: $1v = v$, for all $v \in V$.
(iv) Two distributive laws hold:

$$(a + b)v = av + bv \quad \text{and} \quad a(v + w) = av + aw,$$

for all $a, b \in F$ and $v, w \in V$.

All of Section 1 can be repeated, replacing the field $\mathbb{R}$ by F. Thus the space F^n of row vectors $(a_1, \ldots, a_n)$, $a_i \in F$, is a vector space over F and so on.

It is important to note that the definition of vector space includes implicitly the choice of a field F. The elements of this field F are often called *scalars*. We usually keep this field fixed. Of course, if V is a complex vector space, meaning a vector space over the field $\mathbb{C}$, and if $F \subset \mathbb{C}$ is any subfield, then V is also naturally a vector space over F because cv is defined for all $c \in F$. But we consider the vector space structure to have changed when we restrict the scalars from $\mathbb{C}$ to F.

Two important concepts analogous to subgroups and isomorphisms of groups are the concepts of subspace and of isomorphism of vector spaces. We have already defined subspaces for complex vector spaces, and the definition is the same for any field. A *subspace* W of a vector space V (over a field F) is a subset with the following properties:

(2.12)

(a) If $w, w' \in W$, then $w + w' \in W$.
(b) If $w \in W$ and $c \in F$, then $cw \in W$.
(c) $0 \in W$.

A subspace W is called a *proper* subspace of V if it is neither the whole space V nor the zero subspace $\{0\}$.

It is easy to see that a subspace is just a subset on which the laws of composition induce the structure of vector space.

As in Section 1, the space of all solutions of a system of m linear equations in n unknowns

$$AX = 0,$$

with coefficients in F, is an example of a subspace of the space F^n.

(2.13) **Definition.** An *isomorphism* φ from a vector space V to a vector space V', both over the same field F, is a bijective map $\varphi: V \longrightarrow V'$ compatible with the laws of composition, that is, a bijective map satisfying

$$\text{(a) } \varphi(v + v') = \varphi(v) + \varphi(v') \quad \text{and} \quad \text{(b) } \varphi(cv) = c\varphi(v),$$

for all $v, v' \in V$ and all $c \in F$.

(2.14) **Examples.**

(a) The space F^n of n-dimensional row vectors is isomorphic to the space of n-dimensional column vectors.

(b) View the set of complex numbers $\mathbb{C}$ as a real vector space, as in (1.8b). Then the map $\varphi: \mathbb{R}^2 \longrightarrow \mathbb{C}$ sending $(a, b) \rightsquigarrow a + bi$ is an isomorphism.

3. BASES AND DIMENSION

In this section we discuss the terminology used when working with the two operations, addition and scalar multiplication, in an abstractly given vector space. The new concepts are *span, linear independence,* and *basis*.

It will be convenient to work with *ordered* sets of vectors here. The ordering will be unimportant much of the time, but it will enter in an essential way when we make explicit computations. We've been putting curly brackets around unordered sets, so in order to distinguish ordered from unordered sets, let us enclose ordered sets with round brackets. Thus the ordered set (a, b) is considered different from the ordered set (b, a), whereas the unordered sets $\{a, b\}$ and $\{b, a\}$ are considered equal. Repetitions will also be allowed in an ordered set. So (a, a, b) is considered an ordered set, and it is different from (a, b), in contrast to the convention for unordered sets, where $\{a, a, b\}$ would denote the same set as $\{a, b\}$.

Let V be a vector space over a field F, and let $(v_1, \ldots, v_n)$ be an ordered set of elements of V. A *linear combination* of $(v_1, \ldots, v_n)$ is any vector of the form

$$w = c_1v_1 + c_2v_2 + \cdots + c_nv_n, \; c_i \in F. \tag{3.1}$$

For example, suppose that the ordered set consists of the two vectors in $\mathbb{R}^3$ considered in (1.4): $v_1 = (1,0,1)^t$ and $v_2 = (1,2,0)^t$. Then a linear combination will have the form (1.5): $(c_1 + c_2, 2c_2, c_1)^t$. The vector $(3,4,1)^t = v_1 + 2v_2$ is one such linear combination.

A solution X of a system of linear equations written in the matrix form $AX = B$ [Chapter 1 (1.9)] exhibits the column vector B as a linear combination of the columns of the matrix A. The coefficients are the entries of the vector X.

A linear combination of a single vector (v) is just a multiple cv or v .

The set of all vectors w which are linear combinations of $(v_1,\ldots,v_n)$ forms a subspace W of V, called the subspace *spanned* by the set: If w (3.1) and $w' = c_1'v_1 + \cdots + c_n'v_n$ are elements of W, then so is

$$w + w' = (c_1 + c_1')v_1 + \cdots + (c_n + c_n')v_n,$$

and if $a \in F$, then $aw = (ac_1)v_1 + \cdots + (ac_n)v_n$ is in W. So $w + w'$ and aw are in W. Finally, $0 = 0v_1 + \cdots + 0v_n \in W$. This shows that the conditions of (2.12) hold.

The space spanned by a set S will often be denoted by Span S. Clearly, Span S is the smallest subspace of V which contains S. We could also call it the subspace *generated* by S. Note that the order is irrelevant here. The span of S is the same as the span of any reordering of S.

One can also define the span of an infinite set of vectors. We will discuss this in Section 5. In this section, let us assume that our sets are *finite*.

(3.2) **Proposition.** Let S be a set of vectors of V, and let W be a subspace of V. If $S \subset W$, then Span $S \subset W$.

This is obvious, because W is closed under addition and scalar multiplication. If $S \subset W$, then any linear combination of vectors of S is in W too. □

A *linear relation* among vectors $v_1,\ldots,v_n$ is any relation of the form

(3.3) $$c_1v_1 + c_2v_2 + \cdots + c_nv_n = 0,$$

where the coefficients c_i are in F. An ordered set $(v_1,\ldots,v_n)$ of vectors is called *linearly independent* if there is no linear relation among the vectors in the set, except for the trivial one in which all the coefficients c_i are zero. It is useful to state this condition positively:

(3.4) *Let $(v_1,\ldots,v_n)$ be a linearly independent set. Then from the equation $c_1v_1 + \cdots + c_nv_n = 0$, we can conclude that $c_i = 0$ for every $i = 1,\ldots,n$.*

Conversely, if (3.4) holds, then the vectors are linearly independent.

The vectors (1.4) are linearly independent.

Note that a linearly independent set S can not have any repetitions. For if two vectors v_i, v_j of S are equal, then

$$v_i - v_j = 0$$

is a linear relation of the form (3.3), the other coefficients being zero. Also, no vector v_i of a linearly independent family may be zero, because if it is, then $v_i = 0$ is a linear relation.

A set which is not linearly independent is called *linearly dependent*.

If V is the space F^m and if the vectors $(v_1,\dots,v_n)$ are given explicitly, we can decide linear independence by solving a system of homogeneous linear equations. For to say that a linear combination $x_1v_1 + \cdots + x_nv_n$ is zero means that each coordinate is zero, and this leads to m equations in the n unknowns x_i. For example, consider the set of three vectors

$$(3.5) \qquad v_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, v_3 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}.$$

Let A denote the matrix whose columns are these vectors:

$$(3.6) \qquad A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{bmatrix}.$$

A general linear combination of the vectors will have the form $x_1v_1 + x_2v_2 + x_3v_3$. Bringing the scalar coefficients to the other side, we can write this linear combination in the form AX, where $X = (x_1,x_2,x_3)^t$. Since $\det A = 1$, the equation $AX = 0$ has only the trivial solution, and this shows that (v_1, v_2, v_3) is a linearly independent set. On the other hand, if we add an arbitrary fourth vector v_4 to this set, the result will be linearly dependent, because every system of three homogeneous equations in four unknowns has a nontrivial solution [Chapter 1 (2.17)].

Here are some elementary facts about linear independence.

(3.7) **Proposition.**

(a) Any reordering of a linearly independent set is linearly independent.

(b) If $v_1 \in V$ is a nonzero vector, then the set (v_1) is linearly independent.

(c) A set (v_1, v_2) of two vectors is linearly dependent if and only if either $v_1 = 0$, or else v_2 is a multiple of v_1.

Let us verify the third of these assertions: Assume (v_1, v_2) dependent. Let the relation be $c_1v_1 + c_2v_2 = 0$, where c_1, c_2 are not both zero. If $c_2 \neq 0$, we can solve for v_2:

$$v_2 = \frac{-c_1}{c_2}v_1.$$

In this case v_2 is a multiple of v_1. If $c_2 = 0$, then $c_1 \neq 0$ and the equation shows that $v_1 = 0$. Conversely, if $v_2 = cv_1$, then the relation $cv_1 - v_2 = 0$ shows that the set (v_1, v_2) is linearly dependent, and if $v_1 = 0$, then the relation $v_1 + 0v_2 = 0$ shows the same thing. □

A set of vectors $(v_1,\dots, v_n)$ which is linearly independent and which also spans V is called a *basis*. For example, the vectors (1.4) form a basis for the space of solutions of the linear equation (1.3). We will often use a symbol such as $\mathbf{B}$ to denote a basis.

Let $\mathbf{B} = (v_1,\dots, v_n)$ be a basis. Then since $\mathbf{B}$ spans V, every $w \in V$ can be written as a linear combination (3.1). Since $\mathbf{B}$ is linearly independent, this expression is unique.

(3.8) **Proposition.** *The set $\mathbf{B} = (v_1,\dots, v_n)$ is a basis if and only if every vector $w \in V$ can be written in a **unique** way in the form (3.1).*

Proof. Suppose that $\mathbf{B}$ is a basis and that w is written as a linear combination in two ways, say (3.1) and also $w = c_1'v_1 + \cdots + c_1'v_n$. Then

$$0 = w - w = (c_1 - c_1')v_1 + \cdots + (c_n - c_n')v_n.$$

Hence by (3.4) $c_1 - c_1' = 0,\dots, c_n - c_n' = 0$. Thus the two linear combinations are the same. On the other hand, the definition of linear independence for $\mathbf{B}$ can be restated by saying that 0 has only one expression as a linear combination. This proves the converse. □

(3.9) **Example.** Let $V = F^n$ be the space of column vectors, and let e_i denote the column vector with 1 in the ith position and zeros elsewhere. The n vectors e_i form a basis for F^n called the *standard basis*. This basis was introduced before, in Chapter 1, Section 4. We will denote it by $\mathbf{E}$. Every vector $X = (x_1,\dots, x_n)^t$ has the unique expression

$$X = x_1e_1 + \cdots + x_ne_n$$

as a linear combination of $\mathbf{E} = (e_1,\dots, e_n)$.

The set (3.5) is another basis of $\mathbb{R}^3$.

We now discuss the main facts (3.15–3.17) which relate the three notions of span, linear independence, and basis.

(3.10) **Proposition.** *Let L be a linearly independent ordered set in V, and let $v \in V$ be any vector. Then the ordered set $L' = (L, v)$ obtained by adding v to L is linearly independent if and only if v is not in the subspace spanned by L.*

Proof. Say that $L = (v_1, \dots ,v_r)$. If $v \in \text{Span } L$, then $v = c_1v_1 + \cdots + c_rv_r$ for some $c_i \in F$. Hence

$$c_1v_1 + \cdots + c_rv_r + (-1)v = 0$$

is a linear relation among the vectors of L', and the coefficient -1 is not zero. Thus L' is linearly dependent.

Conversely, suppose that L' is linearly dependent, so that there is some linear relation

$$c_1v_1 + \cdots + c_rv_r + bv = 0,$$

in which not all coefficients are zero. Then certainly $b \neq 0$. For, if b were zero, the expression would reduce to

$$c_1v_1 + \cdots + c_rv_r = 0.$$

Since L is assumed to be linearly independent, we could conclude that $c_1 = \cdots = c_r = 0$ too, contrary to hypothesis. Now that we know $b \neq 0$, we can solve for v:

$$v = \frac{-c_1}{b} v_1 + \cdots + \frac{-c_r}{b} v_r.$$

Thus $v \in \text{Span } L$. □

(3.11) **Proposition.** Let S be an ordered set of vectors, let $v \in V$ be any vector, and let $S' = (S,v)$. Then $\text{Span } S = \text{Span } S'$ if and only if $v \in \text{Span } S$.

Proof. By definition, $v \in \text{Span } S'$. So if $v \notin \text{Span } S$, then $\text{Span } S \neq \text{Span } S'$. Conversely, if $v \in \text{Span } S$, then $S' \subset \text{Span } S$; hence $\text{Span } S' \subset \text{Span } S$ (3.2). The fact that $\text{Span } S' \supset \text{Span } S$ is trivial, and so $\text{Span } S' = \text{Span } S$. □

(3.12) **Definition.** A vector space V is called *finite-dimensional* if there is some finite set S which spans V.

For the rest of this section, we assume that our given vector space V is finite-dimensional.

(3.13) **Proposition.** Any finite set S which spans V contains a basis. In particular, any finite-dimensional vector space has a basis.

Proof. Suppose $S = (v_1,\ldots,v_n)$ and that S is not linearly independent. Then there is a linear relation

$$c_1v_1 + \cdots + c_nv_n = 0$$

in which some c_i is not zero, say $c_n \neq 0$. Then we may solve for v_n:

$$v_n = \frac{-c_1}{c_n} v_1 + \cdots + \frac{-c_{n-1}}{c_n} v_{n-1}.$$

This shows that $v_n \in \text{Span}(v_1,\ldots,v_{n-1})$. Putting $v = v_n$ and $S = (v_1,\ldots,v_{n-1})$ in (3.11), we conclude $\text{Span}(v_1,\ldots,v_{n-1}) = \text{Span}(v_1,\ldots,v_n) = V$. So we may eliminate v_n from S. Continuing this way we eventually obtain a family which is linearly independent but still spans V—a basis.

Note. There is a problem with this proof if V is the zero vector space $\{0\}$. For, starting with an arbitrary collection of vectors in V (all of them equal to zero), our procedure will throw them out, one at a time, until there is only one vector $v_1 = 0$ left. And (0) is a linearly dependent set. How can we eliminate it? Of course the zero vector space is not particularly interesting. But it may lurk around, waiting to trip us up. We have to allow the possibility that a vector space which arises in the course of some computation, such as solving a system of homogeneous linear equations, is the zero space. In order to avoid having to make special mention of this case in the future, we adopt the following conventions:

(3.14) (a) The empty set is linearly independent.
(b) The span of the empty set is the zero subspace.

Thus the empty set is a basis for the zero vector space. These conventions allow us to throw out the last vector $v_1 = 0$, and rescue the proof. □

(3.15) **Proposition.** Let V be a finite-dimensional vector space. Any linearly independent set L can be extended by adding elements, to get a basis.

Proof. Let S be a finite set which spans V. If all elements of S are in Span L, then L spans V (3.2) and so it is a basis. If not, choose $v \in S$, which is not in Span L. By (3.10), (L, v) is linearly independent. Continue until you get a basis. □

(3.16) **Proposition.** Let S, L be finite subsets of V. Assume that S spans V and that L is linearly independent. Then S contains at least as many elements as L does.

Proof. To prove this, we write out what a relation of linear dependence on L means in terms of the set S, obtaining a homogeneous system of m linear equations in n unknowns, where $m = |S|$ and $n = |\mathbf{L}|$. Say that $S = (v_1, \ldots, v_m)$ and $L = (w_1, \ldots, w_n)$. We write each vector w_j as a linear combination of S, which we can do because S spans V, say

$$w_j = a_{1j}v_1 + \cdots + a_{mj}v_m = \sum_i a_{ij}v_i.$$

Let $u = c_1w_1 + \cdots + c_nw_n = \Sigma_j c_jw_j$ be a linear combination. Substituting, we obtain

$$u = \sum_{i,j} c_j a_{ij} v_i.$$

The coefficient of v_i in this sum is $\Sigma_j a_{ij}c_j$. If this coefficient is zero for every i, then $u = 0$. So to find a linear relation among the vectors of L, it suffices to solve the system $\Sigma_j a_{ij}x_j = 0$ of m equations in n unknowns. If $m < n$, then this system has a nontrivial solution [see Chapter 1 (2.17)], and therefore L is linearly dependent. □

(3.17) **Proposition.** Two bases $\mathbf{B}_1, \mathbf{B}_2$ of the vector space V have the same number of elements.

Proof. Put $\mathbf{B}_1 = S$, $\mathbf{B}_2 = L$ in (3.16) to get $|\mathbf{B}_1| \geq |\mathbf{B}_2|$. By symmetry, $|\mathbf{B}_2| \geq |\mathbf{B}_1|$. □

(3.18) **Definition.** The *dimension* of a finite-dimensional vector space V is the number of vectors in a basis. The dimension will be denoted by $\dim V$.

(3.19) **Proposition.**

(a) If S spans V, then $|S| \geq \dim V$, and equality holds only if S is a basis.
(b) If L is linearly independent, then $|L| \leq \dim V$, and equality holds only if L is a basis.

Proof. This follows from (3.13) and (3.15). □

(3.20) **Proposition.** If $W \subset V$ is a subspace of a finite-dimensional vector space, then W is finite-dimensional, and $\dim W \leq \dim V$. Moreover, $\dim W = \dim V$ only if $W = V$.

Proof. This will be obvious, once we show that W is finite-dimensional. For, if $W < V$, that is, if W is contained in but not equal to V, then a basis for W will not span V, but it can be extended to a basis of V by (3.15). Hence $\dim W < \dim V$. We now check finite-dimensionality: If some given linearly independent set L in W does not span W, there is a vector $w \in W$ not in Span L, and by Proposition (3.10), (L, w) is linearly independent. So, we can start with the empty set and add elements of W using (3.10), hoping to end up with a basis of W. Now it is obvious that if L is a linearly independent set in W then it is also linearly independent when viewed as a subset of V. Therefore (3.16) tells us that $|L| \leq n = \dim V$. So the process of adding vectors to L must come to an end after at most n steps. When it is impossible to apply (3.10) again, L is a basis of W. This shows that W is finite-dimensional, as required. □

Notes.

(a) The key facts to remember are (3.13), (3.15), and (3.16). The others follow.
(b) This material is not deep. Given the definitions, you could produce a proof of the main result (3.16) in a few days or less, though your first try would probably be clumsy.

One important example of a vector space is obtained from an arbitrary set S by forming linear combinations of elements of S with coefficients in F in a formal way. If $S = (s_1, \dots, s_n)$ is a finite ordered set whose elements are distinct, then this space $V = V(S)$ is the set of all expressions

$$a_1 s_1 + \cdots + a_n s_n, \quad a_i \in F. \tag{3.21}$$

Addition and scalar multiplication are carried out formally, assuming no relations among the elements s_i:

(3.22)

$$(a_1s_1 + \cdots + a_ns_n) + (b_1s_1 + \cdots + b_ns_n) = (a_1 + b_1)s_1 + \cdots + (a_n + b_n)s_n$$

$$c(a_1s_1 + \cdots + a_ns_n) = (ca_1)s_1 + \cdots + (ca_n)s_n.$$

This vector space is isomorphic to F^n, by the correspondence

(3.23) $$(a_1,\ldots,a_n) \rightsquigarrow a_1s_1 + \cdots + a_ns_n.$$

Therefore the elements s_i, interpreted as the linear combinations

$$s_1 = 1s_1 + 0s_2 + \cdots + 0s_n,$$

form a basis which corresponds to the standard basis of F^n under the isomorphism (3.23). Because of this, $V(S)$ is often referred to as *the space with basis S,* or *the space of formal linear combinations of S*. If S is an infinite set, $V(S)$ is defined to be the space of all finite expressions (3.21), where $s_i \in S$ (see Section 5).

Since $V(S)$ is isomorphic to F^n when S contains n elements, there is no compelling logical reason for introducing it. However, in many applications, $V(S)$ has a natural interpretation. For example, if S is a set of ingredients, then a vector v may represent a recipe. Or if S is a set of points in the plane, then v (3.21) can be interpreted as a set of weights at the points of S.

4. COMPUTATION WITH BASES

The purpose of bases in vector spaces is to provide a method of computation, and we are going to learn to use them in this section. We will consider two topics: how to express a vector in terms of a given basis, and how to relate two different bases of the same vector space.

Suppose we are given a basis $(v_1,\ldots,v_n)$ of a vector space V. Remember: This means that every vector $v \in V$ can be expressed as a linear combination

(4.1) $$v = x_1v_1 + \cdots + x_nv_n, \quad x_i \in F,$$

in exactly one way. The scalars x_i are called the *coordinates* of v, and the column vector

(4.2) $$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

is called the *coordinate vector* of v, with respect to the basis. We pose the problem of computing this coordinate vector.

The simplest case to understand is that V is the space of column vectors F^n.

Let $\mathbf{B} = (v_1,\ldots,v_n)$ be a basis of F^n. Then each element v_i of our basis is a column vector, and so the array $(v_1,\ldots,v_n)$ forms an $n \times n$ matrix. It seems advisable to introduce a new symbol for this matrix, so we will write it as

$$[\mathbf{B}] = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}. \tag{4.3}$$

For example, if $\mathbf{B}$ is the basis

$$v_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 3 \\ 5 \end{bmatrix}, \quad \text{then} \quad [\mathbf{B}] = \begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}. \tag{4.4}$$

If $\mathbf{E} = (e_1,\ldots,e_n)$ is the standard basis, the matrix $[\mathbf{E}]$ is the identity matrix.

A linear combination $x_1v_1 + \cdots + x_nv_n$ can be written as the matrix product

$$[\mathbf{B}]X = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = v_1x_1 + \cdots + v_nx_n, \tag{4.5}$$

where X denotes the column vector $(x_1,\ldots,x_n)^t$. This is another example of block multiplication. The only new feature is that the definition of matrix multiplication has caused the scalar coefficients x_i to migrate to the right side of the vectors, which doesn't matter.

Now if a vector $Y = (y_1,\ldots,y_n)^t$ is given, we can determine its coordinate vector with respect to the basis $\mathbf{B}$ by solving the equation

$$\begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad \text{or} \quad [\mathbf{B}]X = Y \tag{4.6}$$

for the unknown vector X. This is done by inverting the matrix $[\mathbf{B}]$.

(4.7) **Proposition.** Let $\mathbf{B} = (v_1,\ldots,v_n)$ be a basis of F^n, and let $Y \in F^n$ be a vector. The coordinate vector of Y with respect to the basis $\mathbf{B}$ is

$$X = [\mathbf{B}]^{-1}Y. \square$$

Note that we get Y back if $\mathbf{B}$ is the standard basis $\mathbf{E}$, because $[\mathbf{E}]$ is the identity matrix. This is as it should be.

In Example (4.4),

$$[\mathbf{B}]^{-1} = \begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}^{-1} = \begin{bmatrix} -5 & 3 \\ 2 & -1 \end{bmatrix}.$$

So the coordinate vector of $Y = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$ is $X = \begin{bmatrix} 7 \\ -2 \end{bmatrix}$, which means that $Y = 7v_1 - 2v_2$.

Of course we can not solve in this way unless the matrix is invertible. Fortunately, [B] is always invertible, and in fact it can be any invertible matrix.

(4.8) **Proposition.** Let A be an $n \times n$ matrix with entries in a field F. The columns of A form a basis of F^n if and only if A is invertible.

Proof. Denote the ith column of A by v_i. For any column vector $X = (x_1, \dots, x_n)^t$, the matrix product $AX = v_1x_1 + \cdots + v_nx_n$ is a linear combination of the set $(v_1, \dots, v_n)$. So this set is linearly independent if and only if the only solution of the equation $AX = 0$ is the trivial solution $X = 0$. And as we know, this is true if and only if A is invertible [Chapter 1 (2.18)]. Morever, if $(v_1, \dots, v_n)$ is a linearly independent set, then it forms a basis because the dimension of F^n is n. □

Now let V be an abstractly given vector space. We want to use matrix notation to facilitate the manipulation of bases, and the way we have written ordered sets of vectors was chosen with this in mind:

$$(4.9) \qquad (v_1, \dots, v_n).$$

Perhaps this array should be called a *hypervector*. Unless our vectors are given concretely, we won't be able to represent this hypervector by a matrix, so we will work with it formally, as if it were a vector. Since multiplication of two elements of a vector space is not defined, we can not multiply two matrices whose entries are vectors. But there is nothing to prevent us from multiplying the hypervector $(v_1, \dots, v_m)$ by a matrix of scalars. Thus a linear combination of these vectors can be written as the product with a column vector X:

$$(4.10) \qquad (v_1, \dots, v_m)\begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = v_1x_1 + \cdots + v_mx_m.$$

Evaluating the product, we obtain another vector—a linear combination. The scalar coefficients x_i are on the right side of the vectors as before. If we use a symbol such as **B** to denote the set $(v_1, \dots, v_m)$, then the notation for this linear combination becomes very compact: $\mathbf{B}X = v_1x_1 + \cdots + v_nx_n$.

We may also multiply a hypervector on the right by a matrix of scalars. If A is an $m \times n$ matrix, the product will be another hypervector, say $(w_1, \dots, w_n)$:

$$(4.11) \qquad (v_1, \dots, v_m)\begin{bmatrix} & A & \end{bmatrix} = (w_1, \dots, w_n).$$

To evaluate the product, we use the rule for matrix multiplication:

$$(4.12) \qquad w_j = v_1a_{1j} + v_2a_{2j} + \cdots + v_ma_{mj}.$$

So each vector w_j is a linear combination of $(v_1, \dots, v_m)$, and the scalar coefficients in

this linear combination form the columns of the matrix A. That is what the equation means. For example,

$$(v_1, v_2)\begin{bmatrix} 3 & 2 & 1 \\ 4 & 0 & 1 \end{bmatrix} = (3v_1 + 4v_2, 2v_1, v_1 + v_2).$$

Let us restate this formally:

(4.13) **Proposition.** Let $S = (v_1, \dots, v_m)$ and $U = (w_1, \dots, w_n)$ be ordered sets of elements of a vector space V. The elements of U are in the span of S if and only if there is an $m \times n$ scalar matrix A such that $(v_1, \dots, v_m)A = (w_1, \dots, w_n)$. □

Now let us consider the problem of determining the coordinate vector X of a given vector $v \in V$ with respect to a given basis $\mathbf{B} = (v_1, \dots, v_n)$. That is, we wish to write $v = \mathbf{B}X$ explicitly, as in (4.10). It is clear that this is not possible unless both the basis and the vector are given in some explicit way, so we can not solve the problem as posed. But we can use multiplication by the hypervector $\mathbf{B}$ to define abstractly an *isomorphism of vector spaces*

(4.14) $$\psi: F^n \longrightarrow V \text{ sending}$$
$$X \rightsquigarrow \mathbf{B}X,$$

from the space F^n of column vectors to V. This map is bijective because every vector v is a linear combination (4.10) in exactly one way—it is surjective because the set $\mathbf{B}$ spans V, and injective because $\mathbf{B}$ is linearly independent. The axioms for an isomorphism (2.13) are easy to check. We can use this isomorphism to introduce *coordinates* into the vector space V.

The coordinate vector of a vector v is $X = \psi^{-1}(v)$. Please note that the symbol $\mathbf{B}^{-1}$ is not defined. So unless the basis is given more specifically, we won't have an explicit formula for the inverse function ψ^{-1}. But the existence of the isomorphism ψ is of interest in itself:

(4.15) **Corollary.** Every vector space V of dimension n is isomorphic to the space F^n of column vectors. □

Notice that F^n is *not* isomorphic to F^m if $m \neq n$, because F^n has a basis of n elements, and the number of elements in a basis depends only on the vector space, not on the choice of a basis. Thus the finite-dimensional vector spaces V over a field F are completely classified by (4.15): Every V is isomorphic to F^n, for some uniquely determined integer n. It follows that we will know all about an arbitrary vector space if we study the basic examples of column vectors. This reduces any problem on vector spaces to the familiar algebra of column vectors, once a basis is given.

We now come to a very important computational method: *change of basis*. Identifying V with the isomorphic vector space F^n is useful when a natural basis is

presented to us, but not when the given basis is poorly suited to the problem at hand. In that case, we will want to change coordinates. So let us suppose that we are given two bases for the same vector space V, say $\mathbf{B} = (v_1,\dots,v_n)$ and $\mathbf{B}' = (v_1',\dots,v_n')$. We will think of $\mathbf{B}$ as the *old* basis, and $\mathbf{B}'$ as a *new* basis. There are two computations which we wish to clarify. We ask first: How are the two bases related? Secondly, a vector $v \in V$ will have coordinates with respect to each of these bases, but of course they will be different. So we ask: How are the two coordinate vectors related? These are the computations called change of basis. They will be very important in later chapters. They are also confusing and can drive you nuts if you don't organize the notation well.

We begin by noting that since the new basis spans V, every vector of the old basis $\mathbf{B}$ is a linear combination of the new basis $\mathbf{B}' = (v_1',\dots,v_n')$. So Proposition (4.13) tells us that there is an equation of the form

$$(4.16) \qquad (v_1',\dots,v_n')\begin{bmatrix} & P & \end{bmatrix} = (v_1,\dots,v_n), \quad \text{or } \mathbf{B}'P = \mathbf{B},$$

where P is an $n \times n$ matrix of scalars. This matrix equation reads

$$(4.17) \qquad v_1'p_{1j} + v_2'p_{2j} + \cdots + v_n'p_{nj} = v_j,$$

where p_{ij} are the entries of P. The matrix P is called the *matrix of change of basis*. Its jth column is the coordinate vector of the old basis vector v_j, when computed with respect to the new basis $\mathbf{B}'$.

Note that the matrix of change of basis is *invertible*. This can be shown as follows: Interchanging the roles of $\mathbf{B}$ and $\mathbf{B}'$ provides a matrix P' such that $\mathbf{B}P' = \mathbf{B}'$. Combining this with (4.16), we obtain the relation $\mathbf{B}P'P = \mathbf{B}$:

$$(v_1,\dots,v_n)\begin{bmatrix} & P'P & \end{bmatrix} = (v_1,\dots,v_n).$$

This formula expresses each v_i as a linear combination of the vectors $(v_1,\dots v_n)$. The entries of the product matrix $P'P$ are the coefficients. But since $\mathbf{B}$ is a linearly independent set, there is *only one way* to write v_i as such a linear combination of $(v_1,\dots,v_n)$, namely $v_i = v_i$, or $\mathbf{B}I = \mathbf{B}$. So $P'P = I$. This shows that P is invertible.

Now let X be the coordinate vector of v, computed with respect to the old basis $\mathbf{B}$, that is, $v = \mathbf{B}X$. Substituting (4.16) gives us the matrix equation

$$(4.18) \qquad v = \mathbf{B}X = \mathbf{B}'PX.$$

This equation shows that $PX = X'$ is the coordinate vector of v with respect to the new basis $\mathbf{B}'$.

Recapitulating, we have a single matrix P, the matrix of change of basis, with the dual properties

$$(4.19) \qquad \mathbf{B} = \mathbf{B}'P \quad \text{and} \quad PX = X',$$

where X, X' denote the coordinate vectors of an arbitrary vector v with respect to the

two bases. Each of these properties characterizes P. Note the position of the primes carefully.

We can compute the matrix of change of basis explicitly when $V = F^n$ and the old basis is the standard basis $\mathbf{E}$, but where the new basis $\mathbf{B}'$ is arbitrary. The two bases determine matrices $[\mathbf{E}] = I$ and $[\mathbf{B}']$, as in (4.3). Formula (4.19) gives us the matrix equation $I = [\mathbf{B}']P$. Hence the matrix of change of basis is

(4.20) $$P = [\mathbf{B}']^{-1}, \quad \textit{if } V = F^n \textit{ and if the old basis is } \mathbf{E}.$$

We can also write this as $[\mathbf{B}'] = P^{-1}$. So

(4.21) *If the old basis is* $\mathbf{E}$, *the new basis vectors are the columns of* P^{-1}.

In the above discussion, the matrix P was determined in terms of two bases $\mathbf{B}$ and $\mathbf{B}'$. We could also turn the discussion around, starting with just one basis $\mathbf{B}$ and an invertible matrix $P \in GL_n(F)$. Then we can define a new basis by formula (4.16), that is,

(4.22) $$\mathbf{B}' = \mathbf{B}P^{-1}.$$

The vectors v_i making up the old basis are in the span of $\mathbf{B}'$ because $\mathbf{B} = \mathbf{B}'P$ (4.13). Hence $\mathbf{B}'$ spans V and, having the right number of elements, $\mathbf{B}'$ is a basis.

(4.23) **Corollary.** Let $\mathbf{B}$ be a basis of a vector space V. The other bases are the sets of the form $\mathbf{B}' = \mathbf{B}P^{-1}$, where $P \in GL_n(F)$ is an invertible matrix.

It is, of course, unnecessary to put an inverse matrix into this statement. Since P is arbitrary, so is P^{-1}. We could just as well set $P^{-1} = Q$ and say $\mathbf{B}' = \mathbf{B}Q$, where $Q \in GL_n(F)$. □

As an application of our discussion, let us compute the order of the general linear group $GL_2(F)$ when F is the prime field $\mathbb{F}_p$. We do this by computing the number of bases of the vector space $V = F^2$. Since the dimension of V is 2, any linearly independent set (v_1, v_2) of two elements forms a basis. The first vector v_1 of a linearly independent set is not zero. And since the order of F is p, V contains p^2 vectors including 0. So there are $p^2 - 1$ choices for the vector v_1. Next, a set (v_1, v_2) of two vectors, with v_1 nonzero, is linearly independent if and only if v_2 is not a multiple of v_1 (3.7). There are p multiples of a given nonzero vector v_1. Therefore if v_1 is given, there are $p^2 - p$ vectors v_2 such that (v_1, v_2) is linearly independent. This gives us

$$(p^2 - 1)(p^2 - p) = p(p + 1)(p - 1)^2$$

bases for V altogether.

(4.24) **Corollary.** The general linear group $GL_2(\mathbb{F}_p)$ has order $p(p + 1)(p - 1)^2$.

Proof. Proposition (4.23) establishes a bijective correspondence between bases of F^n and elements of $GL_n(F)$. □

5. INFINITE-DIMENSIONAL SPACES

Some vector spaces are too big to be spanned by any finite set of vectors. They are called *infinite-dimensional*. We are not going to need them very often, but since they are so important in analysis, we will discuss them briefly.

The most obvious example of an infinite-dimensional space is the space $\mathbb{R}^\infty$ of infinite real vectors

$$(a) = (a_1, a_2, a_3, \ldots). \tag{5.1}$$

It can also be thought of as the space of sequences $\{a_n\}$ of real numbers. Examples (1.7c, d) are also infinite-dimensional.

The space $\mathbb{R}^\infty$ has many important subspaces. Here are a few examples:

(5.2) **Examples.**

(a) Convergent sequences: $C = \{(a) \in \mathbb{R}^\infty \mid \lim_{n\to\infty} a_n \text{ exists}\}$.

(b) Bounded sequences: $\ell^\infty = \{(a) \in \mathbb{R}^\infty \mid \{a_n\} \text{ is bounded}\}$.

A sequence $\{a_n\}$ is called *bounded* if there is some real number b, a *bound*, such that $|a_n| \leq b$ for all n.

(c) Absolutely convergent series: $\ell^1 = \{(a) \in \mathbb{R}^\infty \mid \sum_1^\infty |a_n| < \infty\}$.

(d) Sequences with finitely many nonzero terms:

$$Z = \{(a) \in \mathbb{R}^\infty \mid a_n = 0 \text{ for all but finitely many } n\}.$$

All of the above subspaces are infinite-dimensional. You should be able to make up some more.

Now suppose that V is a vector space, infinite-dimensional or not. What should we mean by the *span* of an infinite set S of vectors? The difficulty is this: It is not always possible to assign a vector as the value of an infinite linear combination $c_1v_1 + c_2v_2 + \cdots$ in a consistent way. If we are talking about the vector space of real numbers, that is, $v_i \in \mathbb{R}^1$, then a value can be assigned provided that the series $c_1v_1 + c_2v_2 + \cdots$ converges. The same can be done for convergent series of vectors in $\mathbb{R}^n$ or $\mathbb{R}^\infty$. But many series don't converge, and then we don't know what value to assign.

In algebra it is customary to speak only of linear combinations of finitely many vectors. Therefore, the span of an infinite set S must be interpreted as the set of those vectors v which are linear combinations of *finitely many* elements of S:

$$v = c_1v_1 + \cdots + c_rv_r, \quad \text{where} \quad v_1, \ldots, v_r \in S. \tag{5.3}$$

The number r is allowed to be arbitrarily large, depending on the vector v:

$$\operatorname{Span} S = \left\{ \begin{matrix} \text{finite linear combinations} \\ \text{of elements of } S \end{matrix} \right\}. \tag{5.4}$$

With this definition, Propositions (3.2) and (3.11) continue to hold.

For example, let $e_i = (0,\dots,0,1,0,\dots)$ be the vector in $\mathbb{R}^\infty$ with 1 in the ith position as its only nonzero coordinate. Let $S = (e_1, e_2, e_3,\dots)$ be the infinite set of these vectors e_i. The set S does not span $\mathbb{R}^\infty$, because the vector

$$w = (1, 1, 1,\dots)$$

is not a (finite) linear combination. Instead the span of S is the subspace Z (5.2d).

A set S, infinite or not, is called *linearly independent* if there is no *finite* relation

$$c_1 v_1 + \cdots + c_r v_r = 0, \quad v_1,\dots,v_r \in S, \tag{5.5}$$

except for the trivial relation, in which $c_1 = \cdots = c_r = 0$. Again, the number r is allowed to be arbitrary, that is, the condition has to hold for arbitrarily large r and arbitrary vectors $v_1,\dots,v_r \in S$. For example, the set $S' = (w; e_1, e_2, e_3,\dots)$ is linearly independent, if w, e_i are the vectors defined as above. With this definition of linear independence, Proposition (3.10) continues to be true.

As with finite sets, a *basis* S of V is a linearly independent set which spans V. Thus $S = (e_1, e_2,\dots)$ is a basis of the space Z. It can be shown, using the *Axiom of Choice,* that every vector space V has a basis. However, the proof doesn't tell you how to get one. A basis for $\mathbb{R}^\infty$ will have uncountably many elements, and therefore it can not be written down in an explicit way. We won't need bases for infinite-dimensional spaces very often.

Let us go back for a moment to the case that our vector space V is finite-dimensional (3.12), and ask if there can be an *infinite* basis. In Section 3, we saw that any two finite bases have the same number of elements. We will now complete the picture by showing that every basis is finite. The only confusing point is taken care of by the following proposition:

(5.6) **Proposition.** Let V be finite-dimensional, and let S be any set which spans V. Then S contains a finite subset which spans V.

Proof. By assumption, there is some finite set, say $(w_1,\dots,w_m)$, which spans V. Each w_i is a linear combination of finitely many elements of S, since $\operatorname{Span} S = V$. So when we express the vectors $w_1,\dots,w_m$ in terms of the set S, we only need to use finitely many of its elements. The ones we use make up a finite subset $S' \subset S$. So, $(w_1,\dots,w_m) \subset \operatorname{Span} S'$. Since $(w_1,\dots,w_m)$ spans V, so does S'. □

(5.7) **Proposition.** Let V be a finite-dimensional vector space.

(a) Every set S which spans V contains a finite basis.
(b) Every linearly independent set L is finite and therefore extends to a finite basis.
(c) Every basis is finite.

We leave the proof of (5.7) as an exercise. □

6. DIRECT SUMS

Let V be a vector space, and let $W_1,\dots,W_n$ be subspaces of V. Much of the treatment of linear independence and spans of vectors has analogues for subspaces, and we are going to work out these analogues here.

We consider vectors $v \in V$ which can be written as a sum

$$v = w_1 + \cdots + w_n, \tag{6.1}$$

where w_i is a vector in W_i. The set of all such vectors is called the *sum* of the subspaces or their *span*, and is denoted by

$$W_1 + \cdots + W_n = \{v \in V \mid v = w_1 + \cdots + w_n, \quad \text{with } w_i \in W_i\}. \tag{6.2}$$

The sum is a subspace of V, analogous to the span of a set $\{v_1,\dots,v_n\}$ of vectors. Clearly, it is the smallest subspace containing $W_1,\dots,W_n$.

The subspaces $W_1,\dots,W_n$ are called *independent* if no sum $w_1 + \cdots + w_n$ with $w_i \in W$ is zero, except for the trivial sum in which $w_i = 0$ for all i. In other words, the spaces are independent if

$$w_1 + \cdots + w_n = 0 \ \textit{and}\ w_i \in W_i \ \textit{implies}\ w_i = 0 \ \textit{for all}\ i. \tag{6.3}$$

In case the span is the whole space and the subspaces are independent, we say that V is the *direct sum* of $W_1,\dots,W_n$, and we write

$$V = W_1 \oplus \cdots \oplus W_n, \ \textit{if}\ V = W_1 + \cdots + W_n \ \textit{and if}\ W_1,\dots,W_n \ \textit{are independent}. \tag{6.4}$$

This is equivalent to saying that every vector $v \in V$ can be written in the form (6.1) in *exactly one way*.

So, if $W_1,\dots,W_n$ are independent subspaces of a vector space V and if $U = W_1 + \cdots + W_n$ is their sum, then in fact U is their direct sum: $U = W_1 \oplus \cdots \oplus W_n$.

We leave the proof of the following two propositions as an exercise.

(6.5) **Proposition.**

(a) A single subspace W_1 is independent.

(b) Two subspaces W_1, W_2 are independent if and only if $W_1 \cap W_2 = (0)$. □

(6.6) **Proposition.** Let $W_1,\dots,W_n$ be subspaces of a finite-dimensional vector space V, and let $\mathbf{B}_i$ be a basis for W_i.

(a) The ordered set $\mathbf{B}$ obtained by listing the bases $\mathbf{B}_1,\dots,\mathbf{B}_n$ in order is a basis of V if and only if V is the direct sum $W_1 \oplus \cdots \oplus W_n$.

(b) $\dim(W_1 + \cdots + W_n) \le (\dim W_1) + \cdots + (\dim W_n)$, with equality if and only if the spaces are independent. □

(6.7) **Corollary.** Let W be a subspace of a finite-dimensional vector space V. There is another subspace W' such that $V = W \oplus W'$.

Proof. Let $(w_1,\dots,w_d)$ be a basis for W. Extend to a basis $(w_1,\dots,w_d; v_1,\dots,v_{n-d})$ for V (3.15). The span of $(v_1,\dots,v_{n-d})$ is the required subspace W'. □

(6.8) **Example.** Let $v_1,\dots,v_n$ be nonzero vectors, and let W_i be the span of the single vector v_i. This is the one-dimensional subspace which consists of all scalar multiples of v_i: $W_i = \{cv_i\}$. Then $W_1,\dots,W_n$ are independent subspaces if and only if $(v_1,\dots,v_n)$ are independent vectors. This becomes clear if we compare (3.4) and (6.3). The statement in terms of subspaces is actually the neater one, because the scalar coefficients are absorbed.

(6.9) **Proposition.** Let W_1, W_2 be subspaces of a finite-dimensional vector space V. Then

$$\dim W_1 + \dim W_2 = \dim(W_1 \cap W_2) + \dim(W_1 + W_2).$$

Proof. Note first that the intersection of two subspaces is again a subspace. Choose a basis $(u_1,\dots,u_r)$ for the space $W_1 \cap W_2$, where $r = \dim(W_1 \cap W_2)$. This is a linearly independent set, and it is in W_1. Hence we can extend it to a basis of W_1, say

$$(u_1,\dots,u_r; x_1,\dots,x_{m-r}), \tag{6.10}$$

where $m = \dim W_1$. Similarly, we can extend it to a basis

$$(u_1,\dots,u_r; y_1,\dots,y_{n-r}), \tag{6.11}$$

of W_2, where $n = \dim W_2$. The proposition will follow if we show that the set

$$(u_1,\dots,u_r; x_1,\dots,x_{m-r}; y_1,\dots,y_{n-r}) \tag{6.12}$$

is a basis of $W_1 + W_2$.

This assertion has two parts. First, the vectors (6.12) span $W_1 + W_2$. For any vector v in $W_1 + W_2$ is a sum $v = w_1 + w_2$, with $w_i \in W_i$. We can write w_1 as a linear combination of (6.10), and w_2 as a linear combination of (6.11). Collecting terms, we find that v is a linear combination of (6.12).

Next, the vectors (6.11) are linearly independent: Suppose that some linear combination is zero, say

$$a_1u_1 + \cdots + a_ru_r + b_1x_1 + \cdots + b_{m-r}x_{m-r} + c_1y_1 + \cdots + c_{n-r}y_{n-r} = 0.$$

Abbreviate this as $u + x + y = 0$. Solve for y: $y = -u - x \in W_1$. But $y \in W_2$ too. Hence $y \in W_1 \cap W_2$, and so y is a linear combination, say u', of $(u_1,\dots,u_r)$. Then $-u' + y = 0$ is a relation among the vectors (6.11), which are independent. So it must be the trivial relation. This shows that $y = 0$. Thus our original relation reduces to $u + x = 0$. Since (6.10) is a basis, this relation is trivial: $u = 0$ and $x = 0$. So the whole relation was trivial, as required. □

I don't need to learn 8 + 7: I'll remember 8 + 8 and subtract 1.

T. Cuyler Young, Jr.

EXERCISES

1. Real Vector Spaces

1. Which of the following subsets of the vector space of real $n \times n$ matrices is a subspace?
   (a) symmetric matrices ($A = A^t$)
   (b) invertible matrices
   (c) upper triangular matrices
2. Prove that the intersection of two subspaces is a subspace.
3. Prove the cancellation law in a vector space: If $cv = cw$ and $c \neq 0$, then $v = w$.
4. Prove that if w is an element of a subspace W, then $-w \in W$ too.
5. Prove that the classification of subspaces of $\mathbb{R}^3$ stated after (1.2) is complete.
6. Prove that every solution of the equation $2x_1 - x_2 - 2x_3 = 0$ has the form (1.5).
7. What is the description analogous to (1.4) obtained from the particular solutions $u_1 = (2, 2, 1)$ and $u_2 = (0, 2, -1)$?

2. Abstract Fields

1. Prove that the set of numbers of the form $a + b\sqrt{2}$, where a, b are rational numbers, is a field.
2. Which subsets of $\mathbb{C}$ are closed under $+$, $-$, $\times$, and $\div$ but fail to contain 1?
3. Let F be a subset of $\mathbb{C}$ such that F^+ is a subgroup of $\mathbb{C}^+$ and $F^\times$ is a subgroup of $\mathbb{C}^\times$. Prove that F is a subfield of $\mathbb{C}$.
4. Let $V = F^n$ be the space of column vectors. Prove that every subspace W of V is the space of solutions of some system of homogeneous linear equations $AX = 0$.
5. Prove that a nonempty subset W of a vector space satisfies the conditions (2.12) for a subspace if and only if it is closed under addition and scalar multiplication.
6. Show that in Definition (2.3), axiom (ii) can be replaced by the following axiom: $F^\times$ is an abelian group, and $1 \neq 0$. What if the condition $1 \neq 0$ is omitted?
7. Define homomorphism of fields, and prove that every homomorphism of fields is injective.
8. Find the inverse of 5 (modulo p) for $p = 2, 3, 7, 11, 13$.
9. Compute the polynomial $(x^2 + 3x + 1)(x^3 + 4x^2 + 2x + 2)$ when the coefficients are regarded as elements of the fields **(a)** $\mathbb{F}_5$ **(b)** $\mathbb{F}_7$.
10. Consider the system of linear equations $\begin{bmatrix} 8 & 3 \\ 2 & 6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$.
    (a) Solve it in $\mathbb{F}_p$ when $p = 5, 11, 17$.
    (b) Determine the number of solutions when $p = 7$.

11. Find all primes p such that the matrix

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 3 & -1 \\ -2 & 0 & 2 \end{bmatrix}$$

is invertible, when its entries are considered to be in $\mathbb{F}_p$.

12. Solve completely the systems of linear equations $AX = B$, where

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

(a) in $\mathbb{Q}$ **(b)** in $\mathbb{F}_2$ **(c)** in $\mathbb{F}_3$ **(d)** in $\mathbb{F}_7$.

13. Let p be a prime integer. The nonzero elements of $\mathbb{F}_p$ form a group $\mathbb{F}_p^\times$ of order $p - 1$. It is a fact that this group is always cyclic. Verify this for all primes $p < 20$ by exhibiting a generator.

14. **(a)** Let p be a prime. Use the fact that $\mathbb{F}_p^\times$ is a group to prove that $a^{p-1} \equiv 1$ (modulo p) for every integer a not congruent to zero.

(b) Prove *Fermat's Theorem:* For every integer a,

$$a^p \equiv a \text{ (modulo } p).$$

15. **(a)** By pairing elements with their inverses, prove that the product of all nonzero elements of $\mathbb{F}_p$ is -1.

(b) Let p be a prime integer. Prove *Wilson's Theorem:*

$$(p - 1)! \equiv -1 \text{ (modulo } p).$$

16. Consider a system $AX = B$ of n linear equations in n unknowns, where A and B have integer entries. Prove or disprove: If the system has an integer solution, then it has a solution in $\mathbb{F}_p$ for all p.

17. Interpreting matrix entries in the field $\mathbb{F}_2$, prove that the four matrices $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ form a field.

18. The proof of Lemma (2.8) contains a more direct proof of (2.6). Extract it.

3. Bases and Dimension

1. Find a basis for the subspace of $\mathbb{R}^4$ spanned by the vectors $(1, 2, -1, 0)$, $(4, 8, -4, -3)$, $(0, 1, 3, 4)$, $(2, 5, 1, 4)$.

2. Let $W \subset \mathbb{R}^4$ be the space of solutions of the system of linear equations $AX = 0$, where $A = \begin{bmatrix} 2 & 1 & 2 & 3 \\ 1 & 1 & 3 & 0 \end{bmatrix}$. Find a basis for W.

3. **(a)** Show that a subset of a linearly independent set is linearly independent.

(b) Show that any reordering of a basis is also a basis.

4. Let V be a vector space of dimension n over F, and let $0 \le r \le n$. Prove that V contains a subspace of dimension r.

5. Find a basis for the space of symmetric $n \times n$ matrices.
6. Prove that a square matrix A is invertible if and only if its columns are linearly independent.
7. Let V be the vector space of functions on the interval $[0, 1]$. Prove that the functions x^3, $\sin x$, and $\cos x$ are linearly independent.
8. Let A be an $m \times n$ matrix, and let A' be the result of a sequence of elementary row operations on A. Prove that the rows of A span the same subspace as the rows of A'.
9. Let V be a complex vector space of dimension n. Prove that V has dimension $2n$ as real vector space.
10. A complex $n \times n$ matrix is called *hermitian* if $a_{ij} = \overline{a}_{ji}$ for all i, j. Show that the hermitian matrices form a real vector space, find a basis for that space, and determine its dimension.
11. How many elements are there in the vector space $\mathbb{F}_p{}^n$?
12. Let $F = \mathbb{F}_2$. Find all bases of F^2.
13. *Let $F = \mathbb{F}_5$. How many subspaces of each dimension does the space F^3 contain?*
14. (a) Let V be a vector space of dimension 3 over the field $\mathbb{F}_p$. How many subspaces of each dimension does V have?
    (b) Answer the same question for a vector space of dimension 4.
15. (a) Let $F = \mathbb{F}_2$. Prove that the group $GL_2(F)$ is isomorphic to the symmetric group S_3.
    (b) Let $F = \mathbb{F}_3$. Determine the orders of $GL_2(F)$ and of $SL_2(F)$.
16. Let W be a subspace of V.
    (a) Prove that there is a subspace U of V such that $U + W = V$ and $U \cap W = 0$.
    (b) Prove that there is no subspace U such that $W \cap U = 0$ and that $\dim W + \dim U > \dim V$.

4. *Computation with Bases*

1. Compute the matrix P of change of basis in F^2 relating the standard basis $\mathbf{E}$ to $\mathbf{B}' = (v_1, v_2)$, where $v_1 = (1, 3)^t$, $v_2 = (2, 2)^t$.
2. Determine the matrix of change of basis, when the old basis is the standard basis $(e_1, \dots, e_n)$ and the new basis is $(e_n, e_{n-1}, \dots, e_1)$.
3. Determine the matrix P of change of basis when the old basis is (e_1, e_2) and the new basis is $(e_1 + e_2, e_1 - e_2)$.
4. Consider the equilateral coordinate system for $\mathbb{R}^2$, given by the basis $\mathbf{B}'$ in which $v_1 = e_1$ and v_2 is a vector of unit length making an angle of $120°$ with v_1. Find the matrix relating the standard basis $\mathbf{E}$ to $\mathbf{B}'$.
5. (i) Prove that the set $\mathbf{B} = ((1, 2, 0)^t, (2, 1, 2)^t, (3, 1, 1)^t)$ is a basis of $\mathbb{R}^3$.
   (ii) Find the coordinate vector of the vector $v = (1, 2, 3)^t$ with respect to this basis.
   (iii) Let $\mathbf{B}' = ((0, 1, 0)^t, (1, 0, 1)^t, (2, 1, 0)^t)$. Find the matrix P relating $\mathbf{B}$ to $\mathbf{B}'$.
   (iv) For which primes p is $\mathbf{B}$ a basis of $\mathbb{F}_p{}^3$?
6. Let $\mathbf{B}$ and $\mathbf{B}'$ be two bases of the vector space F^n. Prove that the matrix of change of basis is $P = [\mathbf{B}']^{-1}[\mathbf{B}]$.
7. Let $\mathbf{B} = (v_1, \dots, v_n)$ be a basis of a vector space V. Prove that one can get from $\mathbf{B}$ to any other basis $\mathbf{B}'$ by a finite sequence of steps of the following types:

(i) Replace v_i by $v_i + av_j$, $i \neq j$, for some $a \in F$.
(ii) Replace v_i by cv_i for some $c \neq 0$.
(iii) Interchange v_i and v_j.

8. Rewrite the proof of Proposition (3.16) using the notation of Proposition (4.13).

9. Let $V = F^n$. Establish a bijective correspondence between the sets $\mathcal{B}$ of bases of V and $GL_n(F)$.

10. Let F be a field containing 81 elements, and let V be a vector space of dimension 3 over F. Determine the number of one-dimensional subspaces of V.

11. Let $F = \mathbb{F}_p$.
(a) Compute the order of $SL_2(F)$.
(b) Compute the number of bases of F^n, and the orders of $GL_n(F)$ and $SL_n(F)$.

12. **(a)** Let A be an $m \times n$ matrix with $m < n$. Prove that A has no left inverse by comparing A to the square $n \times n$ matrix obtained by adding $(n - m)$ rows of zeros at the bottom.
(b) Let $\mathbf{B} = (v_1, \ldots, v_m)$ and $\mathbf{B}' = (v_1', \ldots, v_n')$ be two bases of a vector space V. Prove that $m = n$ by defining matrices of change of basis and showing that they are invertible.

5. Infinite-Dimensional Spaces

1. Prove that the set $(w; e_1, e_2, \ldots)$ introduced in the text is linearly independent, and describe its span.

2. We could also consider the space of doubly infinite sequences $(a) = (\ldots, a_{-1}, a_0, a_1, \ldots)$, with $a_i \in \mathbb{R}$. Prove that this space is isomorphic to $\mathbb{R}^\infty$.

3. Prove that the space Z is isomorphic to the space of real polynomials.

4. Describe five more infinite-dimensional subspaces of the space $\mathbb{R}^\infty$.

5. For every positive integer, we can define the space ℓ^p to be the space of sequences such that $\Sigma |a_i|^p < \infty$.
(a) Prove that ℓ^p is a subspace of $\mathbb{R}^\infty$.
(b) Prove that $\ell^p < \ell^{p+1}$.

6. Let V be a vector space which is spanned by a countably infinite set. Prove that every linearly independent subset of V is finite or countably infinite.

7. Prove Proposition (5.7).

6. Direct Sums

1. Prove that the space $\mathbb{R}^{n \times n}$ of all $n \times n$ real matrices is the direct sum of the spaces of symmetric matrices ($A = A^t$) and of skew-symmetric matrices ($A = -A^t$).

2. Let W be the space of $n \times n$ matrices whose trace is zero. Find a subspace W' so that $\mathbb{R}^{n \times n} = W \oplus W'$.

3. Prove that the sum of subspaces is a subspace.

4. Prove Proposition (6.5).

5. Prove Proposition (6.6).

Miscellaneous Problems

1. (a) Prove that the set of symbols $\{a + bi \mid a, b \in \mathbb{F}_3\}$ forms a field with nine elements, if the laws of composition are made to mimic addition and multiplication of complex numbers.

(b) Will the same method work for $\mathbb{F}_5$? For $\mathbb{F}_7$? Explain.

***2.** Let V be a vector space over an infinite field F. Prove that V is not the union of finitely many proper subspaces.

***3.** Let W_1, W_2 be subspaces of a vector space V. The formula $\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2)$ is analogous to the formula $|S_1 \cup S_2| = |S_1| + |S_2| - |S_1 \cap S_2|$, which holds for sets. If three sets are given, then

$$|S_1 \cup S_2 \cup S_3| = |S_1| + |S_2| + |S_3| - |S_1 \cap S_2| - |S_1 \cap S_3| - |S_2 \cap S_3| + |S_1 \cap S_2 \cap S_3|.$$

Does the corresponding formula for dimensions of subspaces hold?

4. Let F be a field which is not of characteristic 2, and let $x^2 + bx + c = 0$ be a quadratic equation with coefficients in F. Assume that the discriminant $b^2 - 4c$ is a square in F, that is, that there is an element $\delta \in F$ such that $\delta^2 = b^2 - 4c$. Prove that the quadratic formula $x = (-b + \delta)/2a$ solves the quadratic equation in F, and that if the discriminant is not a square the polynomial has no root in F.

5. (a) What are the orders of the elements $\begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}, \begin{bmatrix} 2 & \\ & 1 \end{bmatrix}$ of $GL_2(\mathbb{R})$?

(b) Interpret the entries of these matrices as elements of $\mathbb{F}_7$, and compute their orders in the group $GL_2(\mathbb{F}_7)$.

6. Consider the function det: $F^{n \times n} \longrightarrow F$, where $F = \mathbb{F}_p$ is a finite field with p elements and $F^{n \times n}$ is the set of 2×2 matrices.

(a) Show that this map is surjective.

(b) Prove that all nonzero values of the determinant are taken on the same number of times.

7. Let A be an $n \times n$ real matrix. Prove that there is a polynomial $f(t) = a_r t^r + a_{r-1} t^{r-1} + \cdots + a_1 t + a_0$ which has A as root, that is, such that $a_r A^r + a_{r-1} A^{r-1} + \cdots + a_1 A + a_0 I = 0$. Do this by showing that the matrices $I, A, A^2, \ldots$ are linearly dependent.

***8.** An *algebraic curve* in $\mathbb{R}^2$ is the locus of zeros of a polynomial $f(x, y)$ in two variables. By a *polynomial path* in $\mathbb{R}^2$, we mean a parametrized path $x = x(t)$, $y = y(t)$, where $x(t)$, $y(t)$ are polynomials in t.

(a) Prove that every polynomial path lies on a real algebraic curve by showing that, for sufficiently large n, the functions $x(t)^i y(t)^j$, $0 \le i, j \le n$, are linearly dependent.

(b) Determine the algebraic curve which is the image of the path $x = t^2 + t$, $y = t^3$ explicitly, and draw it.

Chapter 4

Linear Transformations

That confusions of thought and errors of reasoning
still darken the beginnings of Algebra,
is the earnest and just complaint of sober and thoughtful men.

Sir William Rowan Hamilton

1. THE DIMENSION FORMULA

The analogue for vector spaces of a homomorphism of groups is a map

$$T: V \longrightarrow W$$

from one vector space over a field F to another, which is compatible with addition and scalar multiplication:

$$(1.1) \qquad T(v_1 + v_2) = T(v_1) + T(v_2) \quad \text{and} \quad T(cv) = cT(v),$$

for all v_1, v_2 in V and all $c \in F$. It is customary to call such a map a *linear transformation*, rather than a homomorphism. However, use of the word *homomorphism* would be correct too. Note that a linear transformation is compatible with linear combinations:

$$(1.2) \qquad T\left(\sum_i c_i v_i\right) = \sum_i c_i T(v_i).$$

This follows from (1.1) by induction. Note also that the first of the conditions of (1.1) says that T is a homomorphism of additive groups $V^+ \longrightarrow W^+$.

We already know one important example of a linear transformation, which is in fact the main example: left multiplication by a matrix. Let A be an $m \times n$ matrix with entries in F, and consider A as an operator on column vectors. It defines a linear transformation

$$(1.3) \qquad \begin{aligned} F^n &\xrightarrow{\text{left mult. by } A} F^m \\ X &\rightsquigarrow AX. \end{aligned}$$

Indeed, $A(X_1 + X_2) = AX_1 + AX_2$, and $A(cX) = cAX$.

Another example: Let P_n be the vector space of real polynomial functions of degree $\leq n$, of the form

$$a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0. \tag{1.4}$$

The derivative $\frac{d}{dx}$ is a linear transformation from P_n to P_{n-1}.

Let $T: V \longrightarrow W$ be any linear transformation. We introduce two subspaces

$$\ker T = \textit{kernel} \text{ of } T = \{v \in V \mid T(v) = 0\} \tag{1.5}$$

$$\text{im } T = \textit{image} \text{ of } T = \{w \in W \mid w = T(v) \text{ for some } v \in V\}.$$

As one may guess from the similar case of group homomorphisms (Chapter 2, Section 4), $\ker T$ is a subspace of V and $\text{im } T$ is a subspace of W.

It is interesting to interpret the kernel and image in the case that T is left multiplication by a matrix A. In that case the kernel T is the set of solutions of the homogeneous linear equation $AX = 0$. The image of T is the set of vectors $B \in F^m$ such that the linear equation $AX = B$ has a solution.

The main result of this section is the *dimension formula*, given in the next theorem.

(1.6) **Theorem** Let $T: V \longrightarrow W$ be a linear transformation, and assume that V is finite-dimensional. Then

$$\dim V = \dim(\ker T) + \dim(\text{im } T).$$

The dimensions of $\text{im } T$ and $\ker T$ are called the *rank* and *nullity* of T, respectively. Thus (1.6) reads

$$\dim V = \text{rank} + \text{nullity}. \tag{1.7}$$

Note the analogy with the formula $|G| = |\ker \varphi| \, |\text{im } \varphi|$ for homomorphisms of groups [Chapter 2 (6.15)].

The *rank* and *nullity* of an $m \times n$ matrix A are defined to be the dimensions of the image and kernel of left multiplication by A. Let us denote the rank by r and the nullity by k. Then k is the dimension of the space of solutions of the equation $AX = 0$. The vectors B such that the linear equation $AX = B$ has a solution form the image, a space whose dimension is r. The sum of these two dimensions is n.

Let B be a vector in the image of multiplication by A, so that the equation $AX = B$ has at least one solution $X = X_0$. Let K denote the space of solutions of the homogeneous equation $AX = 0$, the kernel of multiplication by A. Then the set of solutions of $AX = B$ is the additive coset $X_0 + K$. This restates a familiar fact: Adding any solution of the homogeneous equation $AX = 0$ to a particular solution X_0 of the inhomogeneous equation $AX = B$, we obtain another solution of the inhomogeneous equation.

Suppose that A is a square $n \times n$ matrix. If $\det A \neq 0$, then, as we know, the system of equations $AX = B$ has a unique solution for every B, because A is invert-

ible. In this case, $k = 0$ and $r = n$. On the other hand, if $\det A = 0$ then the space K has dimension $k > 0$. By the dimension formula, $r < n$, which implies that the image is not the whole space F^n. This means that not all equations $AX = B$ have solutions. But those that do have solutions have more than one, because the set of solutions of $AX = B$ is a coset of K.

Proof of Theorem (1.6). Say that $\dim V = n$. Let $(u_1, \ldots, u_k)$ be a basis for the subspace $\ker T$, and extend it to a basis of V [Chapter 3 (3.15)]:

$$(1.8) \qquad (u_1, \ldots, u_k; v_1, \ldots, v_{n-k}).$$

Let $w_i = T(v_i)$ for $i = 1, \ldots, n-k$. If we prove that $(w_1, \ldots, w_{n-k}) = S$ is a basis for $\operatorname{im} T$, then it will follow that $\operatorname{im} T$ has dimension $n - k$. This will prove the theorem.

So we must show that S spans $\operatorname{im} T$ and that it is a linearly independent set. Let $w \in \operatorname{im} T$ be arbitrary. Then $w = T(v)$ for some $v \in V$. We write v in terms of the basis (1.8):

$$v = a_1u_1 + \cdots + a_ku_k + b_1v_1 + \cdots + b_{n-k}v_{n-k},$$

and apply T, noting that $T(u_i) = 0$:

$$w = 0 + \cdots + 0 + b_1w_1 + \cdots + b_{n-k}w_{n-k}.$$

Thus w is in the span of S, and so S spans $\operatorname{im} T$.

Next, suppose a linear relation

$$(1.9) \qquad c_1w_1 + \cdots + c_{n-k}w_{n-k} = 0$$

is given, and consider the linear combination $v = c_1v_1 + \cdots + c_{n-k}v_{n-k}$, where v_i are the vectors (1.8). Applying T to v gives

$$T(v) = c_1w_1 + \cdots + c_{n-k}w_{n-k} = 0.$$

Thus $v \in \ker T$. So we may write v in terms of the basis $(u_1, \ldots, u_k)$ of $\ker T$, say $v = a_1u_1 + \cdots + a_ku_k$. Then

$$-a_1u_1 + \cdots + -a_ku_k + c_1v_1 + \cdots + c_{n-k}v_{n-k} = 0.$$

But (1.8) is a basis. So $-a_1 = 0, \ldots, -a_k = 0$, and $c_1 = 0, \ldots, c_{n-k} = 0$. Therefore the relation (1.9) was trivial. This shows that S is linearly independent and completes the proof.

2. THE MATRIX OF A LINEAR TRANSFORMATION

It is not hard to show that every linear transformation $T: F^n \longrightarrow F^m$ is left multiplication by some $m \times n$ matrix A. To see this, consider the images $T(e_j)$ of the standard basis vectors e_j of F^n. We label the entries of these vectors as follows:

$$(2.1) \qquad T(e_j) = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix},$$

and we form the $m \times n$ matrix $A = (a_{ij})$ having these vectors as its columns. We can write an arbitrary vector $X = (x_1, \ldots, x_n)^t$ from F^n in the form $X = e_1x_1 + \cdots + e_nx_n$, putting scalars on the right. Then

$$T(X) = \sum_j T(e_j)x_j = \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} x_1 + \cdots + \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} x_n = AX.$$

For example, the linear transformation $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ such that

$$T(e_1) = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{and} \quad T(e_2) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

is left multiplication by the matrix

$$A = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix}.$$

If $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = e_1x_1 + e_2x_2$, then

$$T(X) = \begin{bmatrix} 1 \\ 2 \end{bmatrix} x_1 + \begin{bmatrix} -1 \\ 0 \end{bmatrix} x_2 = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 \\ 2x_1 \end{bmatrix}.$$

Using the notation established in Section 4 of Chapter 3, we can make a similar computation with an arbitrary linear transformation $T: V \longrightarrow W$, once bases of the two spaces are given. Let $\mathbf{B} = (v_1, \ldots, v_n)$ and $\mathbf{C} = (w_1, \ldots, w_m)$ be bases of V and of W, and let us use the shorthand notation $T(\mathbf{B})$ to denote the hypervector

$$T(\mathbf{B}) = (T(v_1), \ldots, T(v_n)).$$

Since the entries of this hypervector are in the vector space W, and since $\mathbf{C}$ is a basis for that space, there is an $m \times n$ matrix A such that

$$(2.2) \qquad T(\mathbf{B}) = \mathbf{C}A \quad \text{or} \quad (T(v_1), \ldots, T(v_n)) = (w_1, \ldots, w_m) \begin{bmatrix} & A & \end{bmatrix}$$

[Chapter 3 (4.13)]. Remember, this means that for each j,

$$(2.3) \qquad T(v_j) = \sum_i w_i a_{ij} = w_1 a_{1j} + \cdots + w_m a_{mj}.$$

So A is the matrix whose jth column is the coordinate vector of $T(v_j)$. This $m \times n$ matrix $A = (a_{ij})$ is called the *matrix of T with respect to the bases* $\mathbf{B}, \mathbf{C}$. Different choices of the bases lead to different matrices.

In the case that $V = F^n$, $W = F^m$, and the two bases are the standard bases, A is the matrix constructed as in (2.1).

The matrix of a linear transformation can be used to compute the coordinates of the image vector $T(v)$ in terms of the coordinates of v. To do this, we write v in

terms of the basis, say

$$v = \mathbf{B}X = v_1x_1 + \cdots + v_nx_n.$$

Then

$$T(v) = T(v_1)x_1 + \cdots + T(v_n)x_n = T(\mathbf{B})X = \mathbf{C}AX.$$

Therefore the coordinate vector of $T(v)$ is

$$Y = AX,$$

meaning that $T(v) = \mathbf{C}Y$. Recapitulating, the matrix A of the linear transformation has two dual properties:

(2.4) $$T(\mathbf{B}) = \mathbf{C}A \quad \text{and} \quad Y = AX.$$

The relationship between T and A can be explained in terms of the isomorphisms $\psi: F^n \longrightarrow V$ and $\psi': F^m \longrightarrow W$ determined by the two bases [Chapter 3 (4.14)]. If we use ψ and ψ' to identify V and W with F^n and F^m, then T corresponds to left multiplication by A:

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ \uparrow \psi & & \uparrow \psi' \\ F^n & \xrightarrow{\text{mult by } A} & F^m \end{array}, \qquad \begin{array}{ccc} \mathbf{B}X & \rightsquigarrow & \mathbf{C}AX \\ \uparrow & & \uparrow \\ X & \rightsquigarrow & AX \end{array}$$

Going around this square in the two directions gives the same answer: $T \circ \psi = \psi' \circ A$.

Thus any linear transformation between finite-dimensional vector spaces V and W can be identified with matrix multiplication, once bases for the two spaces are chosen. But if we study changes of basis in V and W, we can do much better. Let us ask how the matrix A changes when we make other choices of bases for V and W. Let $\mathbf{B}' = (v_1', \ldots, v_n')$, $\mathbf{C}' = (w_1', \ldots, w_m')$ be new bases for these spaces. We can relate the new basis $\mathbf{B}'$ to the old basis $\mathbf{B}$ by a matrix $P \in GL_n(F)$, as in Chapter 3 (4.19). Similarly, $\mathbf{C}'$ is related to $\mathbf{C}$ by a matrix $Q \in GL_m(F)$. These matrices have the following properties:

(2.6) $$PX = X' \quad \text{and} \quad QY = Y'.$$

Here X and X' denote the coordinate vectors of a vector $v \in V$ with respect to the bases $\mathbf{B}$ and $\mathbf{B}'$, and similarly Y and Y' denote the coordinate vectors of a vector $w \in W$ with respect to $\mathbf{C}$ and $\mathbf{C}'$.

Let A' denote the matrix of T with respect to the new bases, defined as above (2.4), so that $A'X' = Y'$. Then $QAP^{-1}X' = QAX = QY = Y'$. Therefore

(2.7) $$A' = QAP^{-1}.$$

Note that P and Q are arbitrary invertible $n \times n$ and $m \times m$ matrices [Chapter 3 (4.23)]. Hence we obtain the following description of the matrices of a given linear transformation:

(2.8) **Proposition.** Let A be the matrix of a linear transformation T with respect to some given bases $\mathbf{B}, \mathbf{C}$. The matrices A' which represent T with respect to other bases are those of the form

$$A' = QAP^{-1},$$

where $Q \in GL_m(F)$ and $P \in GL_n(F)$ are arbitrary invertible matrices. □

Now given a linear transformation $T: V \longrightarrow W$, it is natural to look for bases $\mathbf{B}, \mathbf{C}$ of V and W such that the matrix of T becomes especially nice. In fact the matrix can be simplified remarkably.

(2.9) **Proposition.**

(a) *Vector space form:* Let $T: V \longrightarrow W$ be a linear transformation. Bases $\mathbf{B}, \mathbf{C}$ can be chosen so that the matrix of T takes the form

$$A = \left[\begin{array}{c|c} I_r & \\ \hline & 0 \end{array}\right], \tag{2.10}$$

where I_r is the $r \times r$ identity matrix, and $r = \operatorname{rank} T$.

(b) *Matrix form:* Given any $m \times n$ matrix A, there are matrices $Q \in GL_m(F)$ and $P \in GL_n(F)$ so that QAP^{-1} has the form (2.10).

It follows from our discussion that these two assertions amount to the same thing. To derive (a) from (b), choose arbitrary bases $\mathbf{B}, \mathbf{C}$ to start with, and let A be the matrix of T with respect to these bases. Applying (b), we can find P, Q so that QAP^{-1} has the required form. Let $\mathbf{B}' = \mathbf{B}P^{-1}$ and $\mathbf{C}' = \mathbf{C}Q^{-1}$ be the new bases, as in Chapter 3 (4.22). Then the matrix of T with respect to the bases $\mathbf{B}', \mathbf{C}'$ is QAP^{-1}. So these new bases are the required ones. Conversely, to derive (b) from (a) we view an arbitrary matrix A as the matrix of the linear transformation "left multiplication by A", with respect to the standard bases. Then (a) and (2.7) guarantee the existence of P, Q so that QAP^{-1} has the required form.

Note that we can interpret QAP^{-1} as the matrix obtained from A by a succession of row and column operations: We write P and Q as products of elementary matrices: $P = E_p \cdots E_1$ and $Q = E_q' \cdots E_1'$ [Chapter 1 (2.18)]. Then $QAP^{-1} = E_q' \cdots E_1' A E_1^{-1} \cdots E_p^{-1}$. Because of the associative law, it does not matter whether the row operations or the column operations are done first. The equation $(E'A)E = E'(AE)$ tells us that row operations commute with column operations.

It is not hard to prove (2.9b) by matrix manipulation, but let us prove (2.9a) using bases instead. Let $(u_1, \ldots, u_k)$ be a basis for $\ker T$. Extend to a basis $\mathbf{B}$ for V: $(v_1, \ldots, v_r; u_1, \ldots, u_k)$, where $r + k = n$. Let $w_i = T(v_i)$. Then, as in the proof of (1.6), $(w_1, \ldots, w_r)$ is a basis for $\operatorname{im} T$. Extend to a basis $\mathbf{C}$ of W: $(w_1, \ldots, w_r; x_1, \ldots, x_s)$. The matrix of T with respect to these bases has the required form. □

Proposition (2.9) is the prototype for a number of results which will be proved later. It shows the power of working in vector spaces without fixed bases (or coordinates), because the structure of an arbitrary linear transformation is related to the very simple matrix (2.10). It also tells us something remarkable about matrix multiplication, because left multiplication by A on F^m is a linear transformation. Namely, it says that left multiplication by A is the same as left multiplication by a matrix of the form (2.10), but with reference to different coordinate systems. Since multiplication by the matrix (2.10) is easy to describe, we have learned something new.

3. LINEAR OPERATORS AND EIGENVECTORS

Let us now consider the case of a linear transformation $T: V \longrightarrow V$ of a vector space to itself. Such a linear transformation is called a *linear operator* on V. Left multiplication by an $n \times n$ matrix with entries in F defines a linear operator on the space F^n of column vectors.

For example, a rotation ρ_θ of the plane through an angle θ is a linear operator on $\mathbb{R}^2$, whose matrix with respect to the standard basis is

$$(3.1) \qquad R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

To verify that this matrix represents a rotation, we write a vector $X \in \mathbb{R}^2$ in polar coordinates, as $X = (r, \alpha)$. Then in rectangular coordinates, $X = \begin{bmatrix} r\cos\alpha \\ r\sin\alpha \end{bmatrix}$. The addition formulas for sine and cosine show that $RX = \begin{bmatrix} r\cos(\alpha + \theta) \\ r\sin(\alpha + \theta) \end{bmatrix}$. So in polar coordinates, $RX = (r, \alpha + \theta)$. This shows that RX is obtained from X by rotation through the angle θ.

The discussion of the previous section must be changed slightly when we are dealing with linear operators. It is clear that we want to pick only one basis $\mathbf{B} = (v_1, \ldots, v_n)$ for V, and use it in place of both of the bases $\mathbf{B}$ and $\mathbf{C}$ considered in Section 2. In other words, we want to write

$$(3.2) \qquad T(\mathbf{B}) = \mathbf{B}A$$

or

$$T(v_j) = \sum_i v_i a_{ij} = v_1 a_{1j} + \cdots + v_n a_{nj}.$$

This defines the matrix $A = (a_{ij})$ of T. It is a square matrix whose jth column is the coordinate vector of $T(v_j)$ with respect to the basis $\mathbf{B}$. Formula (2.4) is unchanged, provided that W and $\mathbf{C}$ are replaced by V and $\mathbf{B}$. As in the previous section, if X and Y denote the coordinate vectors of v and $T(v)$ respectively, then

$$(3.3) \qquad Y = AX.$$

The new feature arises when we study the effect of a change of basis on V. Suppose that $\mathbf{B}$ is replaced by a new basis $\mathbf{B}' = (v_1',\dots,v_n')$. Then formula (2.7) shows that the new matrix A' has the form

$$A' = PAP^{-1}, \tag{3.4}$$

where P is the matrix of change of basis. Thus the rule for change of basis in a linear transformation gets replaced by the following rule:

(3.5) **Proposition.** Let A be the matrix of a linear operator T with respect to a basis $\mathbf{B}$. The matrices A' which represent T for different bases are those of the form

$$A' = PAP^{-1},$$

for arbitrary $P \in GL_n(F)$. □

In general, we say that a square matrix A is *similar* to A' if $A' = PAP^{-1}$ for some $P \in GL_n(F)$. We could also use the word *conjugate* [see Chapter 2 (3.4)].

Now given A, it is natural to ask for a similar matrix A' which is particularly simple. One may hope to get a result somewhat like (2.10). But here our allowable change is much more restricted, because we have only one basis, and therefore one matrix P, to work with.

We can get some insight into the problem by writing the hypothetical matrix P as a product of elementary matrices: $P = E_r \cdots E_1$. Then

$$PAP^{-1} = E_r \cdots E_1 A E_1^{-1} \cdots E_r^{-1}.$$

In terms of elementary operations, we are allowed to change A by a sequence of steps $A \rightsquigarrow EAE^{-1}$. In other words, we may perform an arbitrary row operation E, but then we must also make the inverse column operation E^{-1}. Unfortunately, the row and column operations interfere with each other, and this makes the direct analysis of such operations confusing. I don't know how to use them. It is remarkable that a great deal can be done by another method.

The main tools for analyzing linear operators are the concepts of eigenvector and invariant subspace.

Let $T: V \longrightarrow V$ be a linear operator on a vector space. A subspace W of V is called an *invariant subspace* or a *T-invariant subspace* if it is carried to itself by the operator:

$$TW \subset W. \tag{3.6}$$

In other words, W is T-invariant if $T(w) \in W$ for all $w \in W$. When this is so, T defines a linear operator on W, called the *restriction* of T to W.

Let W be a T-invariant subspace, and let us choose a basis $\mathbf{B}$ of V by appending some vectors to a basis $(w_1,\dots,w_k)$ of W:

$$\mathbf{B} = (w_1,\dots,w_k, v_1,\dots,v_{n-k}).$$

Then the fact that W is invariant can be read off from the matrix M of T. For, the columns of this matrix are the coordinate vectors of the image vectors [see (2.3)],

and $T(w_j)$ is in the subspace W, so it is a linear combination of the basis $(w_1, \ldots, w_k)$. So when we write $T(w_j)$ in terms of the basis $\mathbf{B}$, the coefficients of the vectors $v_1, \ldots, v_{n-k}$ are zero. It follows that M has the block form

$$M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}, \tag{3.7}$$

where A is a $k \times k$ matrix. Moreover, A is the matrix of the restriction of T to W.

Suppose that $V = W_1 \oplus W_2$ is the direct sum of two T-invariant subspaces, and let $\mathbf{B}_i$ be a basis of W_i. Then we can make a basis $\mathbf{B}$ of V by listing the elements of $\mathbf{B}_1$ and $\mathbf{B}_2$ in succession [Chapter 3 (6.6a)]. In this case the matrix of T will have the block diagonal form

$$M = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}, \tag{3.8}$$

where A_i is the matrix of T restricted to W_i.

The concept of an eigenvector is closely related to that of an invariant subspace. An *eigenvector* v for a linear operator T is a nonzero vector such that

$$T(v) = cv \tag{3.9}$$

for some scalar $c \in F$. Here c is allowed to take the value 0, but the vector v can not be zero. Geometrically, if $V = \mathbb{R}^n$, an eigenvector is a nonzero vector v such that v and $T(v)$ are parallel.

The scalar c appearing in (3.9) is called the *eigenvalue* associated to the eigenvector v. When we speak of an *eigenvalue* of a linear operator T, we mean a scalar $c \in F$ which is the eigenvalue associated to some eigenvector.

For example, the standard basis vector e_1 is an eigenvector for left multiplication by the matrix

$$\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}.$$

The eigenvalue associated to the eigenvector e_1 is 3. Or, the vector $(0,1,1)^t$ is an eigenvector for multiplication by the matrix

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 1 & 1 \\ 3 & 0 & 2 \end{bmatrix}$$

on the space $\mathbb{R}^3$ of column vectors, and its eigenvalue is 2.

Sometimes eigenvectors and eigenvalues are called *characteristic vectors* and *characteristic values*.

Let v be an eigenvector for a linear operator T. The subspace W spanned by v is T-invariant, because $T(av) = acv \in W$ for all $a \in F$. Conversely, if this subspace is invariant, then v is an eigenvector. So an eigenvector can be described as a basis

of a one-dimensional T-invariant subspace. If v is an eigenvector, and if we extend it to a basis $(v = v_1, \dots, v_n)$ of V, then the matrix of T will have the block form

$$\begin{bmatrix} c & B \\ 0 & D \end{bmatrix} = \left[\begin{array}{c|ccc} c & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & * & \\ 0 & & & \end{array}\right],$$

where c is the eigenvalue associated to v_1. This is the block decomposition (3.7) in the case of an invariant subspace of dimension 1.

When we speak of an *eigenvector* for an $n \times n$ matrix A, we mean a vector which is an eigenvector for left multiplication by A, a nonzero column vector such that

$$AX = cX, \quad \text{for some } c \in F.$$

As before, the scalar c is called an *eigenvalue*. Suppose that A is the matrix of T with respect to a basis $\mathbf{B}$, and let X denote the coordinate vector of a vector $v \in V$. Then $T(v)$ has coordinates AX (2.4). Hence X is an eigenvector for A if and only if v is an eigenvector for T. Moreover, if so, then the eigenvalues are the same: T and A have the same eigenvalues.

(3.10) **Corollary.** Similar matrices have the same eigenvalues.

This follows from the fact (3.5) that similar matrices represent the same linear transformation. □

Eigenvectors aren't always easy to find, but it is easy to tell whether or not a given vector X is an eigenvector for a matrix A. We need only check whether or not AX is a multiple of X. So we can tell whether or not a given vector v is an eigenvector for a linear operator T, provided that the coordinate vector of v and the matrix of T with respect to a basis are known. If we do this for one of the basis vectors, we find the following criterion:

(3.11) *The basis vector v_j is an eigenvector of T, with eigenvalue c, if and only if the jth column of A has the form ce_j.*

For the matrix A is defined by the property $T(v_j) = v_1 a_{1j} + \cdots + v_n a_{nj}$. So if $T(v_j) = cv_j$, then $a_{jj} = c$ and $a_{ij} = 0$ if $i \neq j$. □

(3.12) **Corollary.** With the above notation, A is a diagonal matrix if and only if every basis vector v_j is an eigenvector. □

(3.13) **Corollary.** The matrix A of a linear transformation is similar to a diagonal matrix if and only if there is a basis $\mathbf{B}' = (v_1', \dots, v_n')$ of V made up of eigenvectors. □

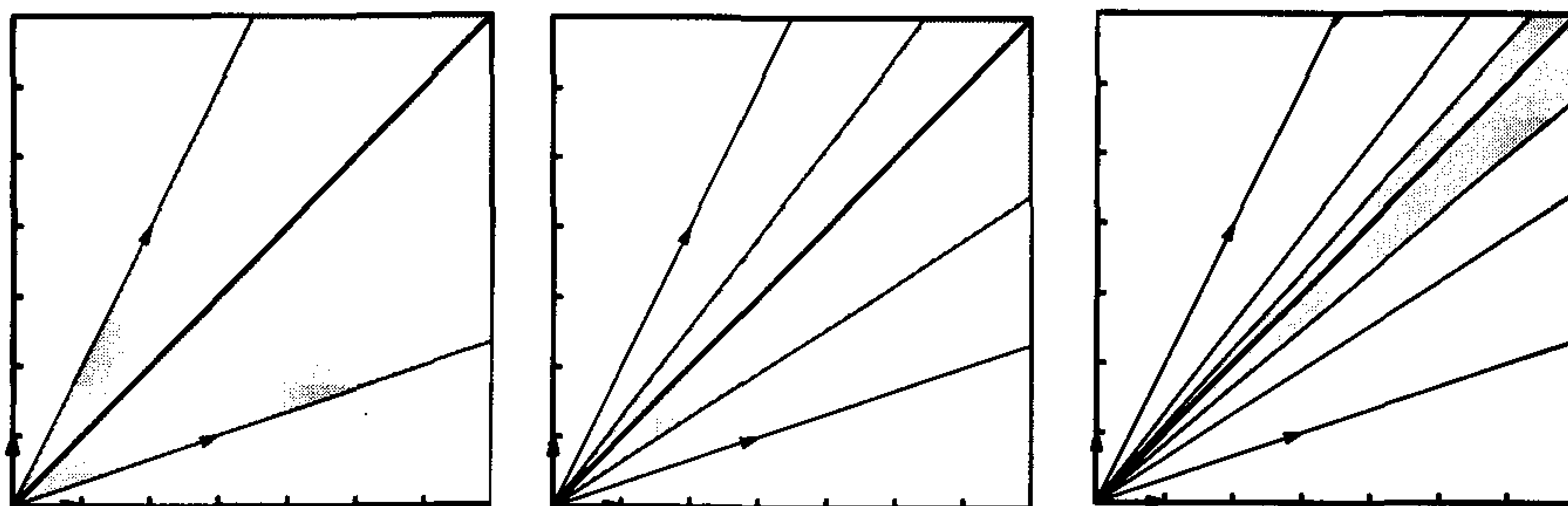
This last corollary shows that we can represent a linear operator very simply by a diagonal matrix, provided that it has enough eigenvectors. We will see in Section 4 that every linear operator on a *complex* vector space has at least one eigenvector, and in Section 6 that in most cases the eigenvectors form a basis. But a linear operator on a real vector space needn't have an eigenvector. For example, the rotation ρ_θ (3.1) of the plane does not carry any vector to a parallel one, unless $\theta = 0$ or π. So ρ_θ has no eigenvector unless $\theta = 0$ or π.

The situation is quite different for real matrices having positive entries. Such matrices are sometimes called *positive* matrices. They occur often in applications, and one of their most important properties is that they always have an eigenvector whose coordinates are positive (a *positive* eigenvector). Instead of proving this fact, let us illustrate it in the case of two variables by examining the effect of multiplication by a positive 2×2 matrix A on $\mathbb{R}^2$.

Let $w_i = Ae_i$. The parallelogram law for vector addition shows that A sends the first quadrant S to the sector bounded by the vectors w_1, w_2. And the coordinate vector of w_i is the ith column of A. Since the entries of A are positive, the vectors w_i lie in the first quadrant. So A carries the first quadrant to itself: $S \supset AS$. Applying A again, we find $AS \supset A^2S$, and so on:

(3.14) $$S \supset AS \supset A^2S \supset A^3S \supset \ldots,$$

as illustrated below in Figure (3.15) for the matrix $A = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$.



(3.15) **Figure.** Images of the first quadrant under repeated multiplication by a positive matrix.

Now the intersection of a nested set of sectors is either a sector or a half line. In our case, the intersection $Z = \bigcap A^rS$ turns out to be a half line. This is intuitively plausible, and it can be shown in various ways. The proof is left as an exercise. We multiply the relation $Z = \bigcap A^rS$ on both sides by A:

$$AZ = A\left(\bigcap_0^\infty A^rS\right) = \bigcap_1^\infty A^rS = Z.$$

Hence $Z = AZ$. This shows that the nonzero vectors in Z are eigenvectors. □

4. THE CHARACTERISTIC POLYNOMIAL

In this section we determine the eigenvectors of an arbitrary linear operator T. Recall that an eigenvector for T is a nonzero vector v such that

$$(4.1) \qquad T(v) = cv,$$

for some c in F. At first glance, it seems difficult to find eigenvectors if the matrix of the linear operator is complicated. The trick is to solve a different problem, namely to determine the *eigenvalues* first. Once an eigenvalue c is determined, equation (4.1) becomes linear in the coordinates of v, and solving it presents no problem.

We begin by writing (4.1) in the form

$$(4.2) \qquad [T - cI](v) = 0,$$

where I stands for the identity operator and $T - cI$ is the linear operator defined by

$$(4.3) \qquad [T - cI](v) = T(v) - cv.$$

It is easy to check that $T - cI$ is indeed a linear operator. If A is the matrix of T with respect to some basis, then the matrix of $T - cI$ is $A - cI$.

We can restate (4.2) as follows:

(4.4) *v is in the kernel of $T - cI$.*

(4.5) **Lemma.** The following conditions on a linear operator $T: V \longrightarrow V$ on a finite-dimensional vector space are equivalent:

(a) $\ker T > 0$.
(b) $\operatorname{im} T < V$.
(c) If A is the matrix of the operator with respect to an arbitrary basis, then $\det A = 0$.
(d) 0 is an eigenvalue of T.

Proof. The dimension formula (1.6) shows that $\ker T > 0$ if and only if $\operatorname{im} T < V$. This is true if and only if T is not an isomorphism, or, equivalently, if and only if A is not an invertible matrix. And we know that the square matrices A which are not invertible are those with determinant zero. This shows the equivalence of (a), (b), and (c). Finally, the nonzero vectors in the kernel of T are the eigenvectors with eigenvalue zero. Hence (a) is equivalent to (d). □

The conditions (4.5a) and (4.5b) are not equivalent for infinite-dimensional vector spaces. For example, let $V = \mathbb{R}^\infty$ be the space of infinite row vectors $(a_1, a_2, \ldots)$, as in Section 5 of Chapter 3. The *shift operator*, defined by

$$(4.6) \qquad T(a_1, a_2, \ldots) = (0, a_1, a_2, \ldots),$$

is a linear operator on V. For this operator, $\ker T = 0$ but $\operatorname{im} T < V$.

(4.7) **Definition.** A linear operator T on a finite-dimensional vector space V is called *singular* if it satisfies any of the equivalent conditions of (4.5). Otherwise, T is *nonsingular*.

We know that c is an eigenvalue for the operator T if and only if $T - cI$ has a nonzero kernel (4.4). So, if we replace T by $T - cI$ in the lemma above, we find:

(4.8) **Corollary.** The eigenvalues of a linear operator T are the scalars $c \in F$ such that $T - cI$ is singular. □

If A is the matrix of T with respect to some basis, then the matrix of $T - cI$ is $A - cI$. So $T - cI$ is singular if and only if $\det(A - cI) = 0$. This determinant can be computed explicitly, and doing so provides us with a concrete method for determining the eigenvalues and eigenvectors.

Suppose for example that A is the matrix

$$\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} \tag{4.9}$$

whose action on $\mathbb{R}^2$ is illustrated in Figure (3.15). Then

$$A - cI = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} - \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix} = \begin{bmatrix} 3-c & 2 \\ 1 & 4-c \end{bmatrix}$$

and

$$\det(A - cI) = c^2 - 7c + 10 = (c-5)(c-2).$$

This determinant vanishes if $c = 5$ or 2, so we have shown that the eigenvalues of A are 5 and 2. To find the eigenvectors, we solve the two systems of linear equations $[A - 5I]X = 0$ and $[A - 2I]X = 0$. The solutions are unique up to scalar factor:

$$v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}. \tag{4.10}$$

Note that the eigenvector v_1 with eigenvalue 5 is in the first quadrant. It lies on the half line Z which is illustrated in Figure (3.15).

We now make the same computation with an arbitrary matrix. It is convenient to change sign. Obviously $\det(cI - A) = 0$ if and only if $\det(A - cI) = 0$. Also, it is customary to replace the symbol c by a variable t. We form the matrix $tI - A$:

$$tI - A = \begin{bmatrix} (t-a_{11}) & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & (t-a_{22}) & \cdots & -a_{2n} \\ \vdots & & & \vdots \\ -a_{n1} & \cdots & \cdots & (t-a_{nn}) \end{bmatrix}. \tag{4.11}$$

Then the complete expansion of the determinant [Chapter 1 (4.11)] shows that $\det(tI - A)$ is a polynomial of degree n in t, whose coefficients are scalars.

(4.12) **Definition.** The *characteristic polynomial* of a linear operator T is the polynomial

$$p(t) = \det(tI - A),$$

where A is the matrix of T with respect to some basis.

The eigenvalues of T are determined by combining (4.8) and (4.12): c is an eigenvalue if and only if $p(c) = 0$.

(4.13) **Corollary.** The eigenvalues of a linear operator are the roots of its characteristic polynomial. □

(4.14) **Corollary.** The eigenvalues of an upper or lower triangular matrix are its diagonal entries.

Proof. If A is an upper triangular matrix, then so is $tI - A$. The determinant of a triangular matrix is the product of its diagonal entries, and the diagonal entries of $tI - A$ are $t - a_{ii}$. Therefore the characteristic polynomial is $p(t) = (t - a_{11})(t - a_{22}) \cdots (t - a_{nn})$, and its roots, the eigenvalues, are $a_{11}, \ldots, a_{nn}$. □

We can compute the characteristic polynomial of an arbitrary 2×2 matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

without difficulty. It is

$$(4.15) \qquad \det(tI - A) = \det\begin{bmatrix} t-a & -b \\ -c & t-d \end{bmatrix} = t^2 - (a + d)t + (ad - bc).$$

The discriminant of this polynomial is

$$(4.16) \qquad (a + d)^2 - 4(ad - bc) = (a - d)^2 + 4bc.$$

If the entries of A are positive real numbers, then the discriminant is also positive, and therefore the characteristic polynomial has real roots, as predicted at the end of Section 3.

(4.17) **Proposition.** The characteristic polynomial of an operator T does not depend on the choice of a basis.

Proof. A second basis leads to a matrix $A' = PAP^{-1}$ [see (3.4)]. We have

$$tI - A' = tI - PAP^{-1} = P(tI)P^{-1} - PAP^{-1} = P(tI - A)P^{-1}.$$

Thus

$$\det(tI - A') = \det(P(tI - A)P^{-1}) = \det P \det(tI - A)\det P^{-1} = \det(tI - A).$$

So the characteristic polynomials computed with A and A' are equal, as was asserted. □

(4.18) **Proposition.** The characteristic polynomial $p(t)$ has the form

$$p(t) = t^n - (\operatorname{tr} A)t^{n-1} + (\textit{intermediate terms}) + (-1)^n(\det A),$$

where tr A, the *trace* of A, is the sum of the diagonal entries:

$$\operatorname{tr} A = a_{11} + a_{22} + \cdots + a_{nn}.$$

All coefficients are independent of the basis. For instance $\operatorname{tr} PAP^{-1} = \operatorname{tr} A$.

This is proved by computation. The independence of the basis follows from (4.17). □

Since the characteristic polynomial, the trace, and the determinant are independent of the basis, they depend only on the operator T. So we may define the terms *characteristic polynomial, trace,* and *determinant* of a linear operator T to be those obtained using the matrix of T with respect to an arbitrary basis.

(4.19) **Proposition.** Let T be a linear operator on a finite-dimensional vector space V.

(a) If V has dimension n, then T has at most n eigenvalues.
(b) If F is the field of complex numbers and $V \neq 0$, then T has at least one eigenvalue, and hence it has an eigenvector.

Proof.

(a) A polynomial of degree n can have at most n different roots. This is true for any field F, though we have not proved it yet [see Chapter 11, (1.8)]. So we can apply (4.13).
(b) Every polynomial of positive degree with complex coefficients has at least one complex root. This fact is called the Fundamental Theorem of Algebra. There is a proof in Chapter 13 (9.1). □

For example, let A be the rotation (3.1) of the real plane $\mathbb{R}^2$ by an angle θ. Its characteristic polynomial is

$$p(t) = t^2 - (2\cos\theta)t + 1, \tag{4.20}$$

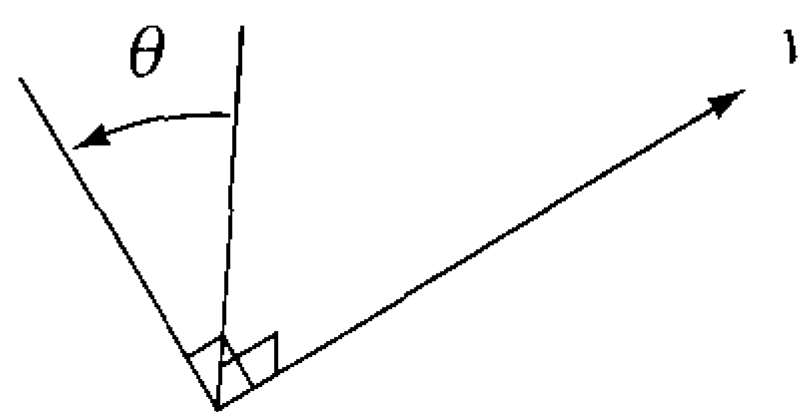
which has no real root unless $\cos\theta = \pm 1$. But if we view A as an operator on $\mathbb{C}^2$, there are two complex eigenvalues.

5. ORTHOGONAL MATRICES AND ROTATIONS

In this section we describe the rotations of two- and three-dimensional spaces $\mathbb{R}^2$ and $\mathbb{R}^3$ about the origin as linear operators. We have already noted (3.1) that a rotation of $\mathbb{R}^2$ through an angle θ is represented as multiplication by the matrix

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

A rotation of $\mathbb{R}^3$ about the origin can be described by a pair (v, θ) consisting of a *unit vector* v, a vector of length 1, which lies in the axis of rotation, and a nonzero angle θ, the angle of rotation. The two pairs (v, θ) and $(-v, -\theta)$ represent the same rotation. We also consider the identity map to be a rotation, though its axis is indeterminate.



(5.1) **Figure.**

The matrix representing a rotation through the angle θ about the vector e_1 is obtained easily from the 2×2 rotation matrix. It is

(5.2) $$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}.$$

Multiplication by A fixes the first coordinate x_1 of a vector and operates by rotation on $(x_2, x_3)^t$. All rotations of $\mathbb{R}^3$ are linear operators, but their matrices can be fairly complicated. The object of this section is to describe these rotation matrices.

A real $n \times n$ matrix A is called *orthogonal* if $A^t = A^{-1}$, or, equivalently, if $A^tA = I$. The orthogonal $n \times n$ matrices form a subgroup of $GL_n(\mathbb{R})$ denoted by O_n and called the *orthogonal group:*

(5.3) $$O_n = \{A \in GL_n(\mathbb{R}) \mid A^tA = I\}.$$

The determinant of an orthogonal matrix is ± 1, because if $A^tA = I$, then

$$(\det A)^2 = (\det A^t)(\det A) = 1.$$

The orthogonal matrices having determinant $+1$ form a subgroup called the *special orthogonal group* and denoted by SO_n:

(5.4) $$SO_n = \{A \in GL_n(\mathbb{R}) \mid A^tA = I, \det A = 1\}.$$

This subgroup has one coset in addition to SO_n, namely the set of elements with determinant -1. So it has index 2 in O_n.

The main fact which we will prove about rotations is stated below:

(5.5) **Theorem.** The rotations of $\mathbb{R}^2$ or $\mathbb{R}^3$ about the origin are the linear operators whose matrices with respect to the standard basis are orthogonal and have determinant 1. In other words, a matrix A represents a rotation of $\mathbb{R}^2$ (or $\mathbb{R}^3$) if and only if $A \in SO_2$ (or SO_3).

Note the following corollary:

(5.6) **Corollary.** The composition of two rotations of $\mathbb{R}^3$ about the origin is also a rotation.

This corollary follows from the theorem because the matrix representing the composition of two linear operators is the product matrix, and because SO_3, being a subgroup of $GL_3(\mathbb{R})$, is closed under products. It is far from obvious geometrically. Clearly, the composition of two rotations about the same axis is also a rotation about that axis. But imagine composing rotations about different axes. What is the axis of rotation of the composed operator?

Because their elements represent rotations, the groups SO_2 and SO_3 are called the two- and three-dimensional *rotation groups*. Things become more complicated in dimension > 3. For example, the matrix

$$\begin{bmatrix} \cos\theta & -\sin\theta & & \\ \sin\theta & \cos\theta & & \\ & & \cos\eta & -\sin\eta \\ & & \sin\eta & \cos\eta \end{bmatrix} \tag{5.7}$$

is an element of SO_4. Left multiplication by this matrix is the composition of a rotation through the angle θ on the first two coordinates and a rotation through the angle η on the last two. Such an operation can not be realized as a single rotation.

The proof of Theorem (5.5) is not very difficult, but it would be clumsy if we did not first introduce some terminology. So we will defer the proof to the end of the section.

To understand the relationship between orthogonal matrices and rotations, we will need the dot product of vectors. By definition, the *dot product* of column vectors X and Y is

$$(X \cdot Y) = x_1y_1 + x_2y_2 + \cdots + x_ny_n. \tag{5.8}$$

It is sometimes useful to write the dot product in matrix form as

$$(X \cdot Y) = X^tY. \tag{5.9}$$

There are two main properties of the dot product of vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$. The first is that $(X \cdot X)$ is the square of the length of the vector:

$$|X|^2 = x_1^2 + x_2^2 \quad \text{or} \quad x_1^2 + x_2^2 + x_3^2,$$

according to the case. This property, which follows from Pythagoras's theorem, is the basis for the definition of length of vectors in $\mathbb{R}^n$: The *length* of X is defined by the formula

$$|X|^2 = (X \cdot X) = x_1^2 + \cdots + x_n^2. \tag{5.10}$$

The *distance* between two vectors X, Y is defined to be the length $|X - Y|$ of $X - Y$.

The second important property of dot product in $\mathbb{R}^2$ and $\mathbb{R}^3$ is the formula

$$(X \cdot Y) = |X|\,|Y| \cos\theta, \tag{5.11}$$

where θ is the angle between the vectors. This formula is a consequence of the law of cosines

$$c^2 = a^2 + b^2 - 2ab \cos \theta$$

for the side lengths a, b, c of a triangle, where θ is the angle subtended by the sides a, b. To derive (5.11), we apply the law of cosines to the triangle with vertices $0, X, Y$. Its side lengths are $|X|$, $|Y|$ and $|X - Y|$, so the law of cosines can be written as

$$(X - Y \cdot X - Y) = (X \cdot X) + (Y \cdot Y) - 2|X|\,|Y| \cos \theta.$$

The left side expands to

$$(X - Y \cdot X - Y) = (X \cdot X) - 2(X \cdot Y) + (Y \cdot Y),$$

and formula (5.11) is obtained by comparing terms.

The most important application of (5.11) is that two vectors X and Y are orthogonal, meaning that the angle θ is $\pi/2$, if and only if $(X \cdot Y) = 0$. This property is taken as the definition of orthogonality of vectors in $\mathbb{R}^n$:

(5.12) $\qquad$ *X is orthogonal to Y if $(X \cdot Y) = 0$.*

(5.13) **Proposition.** The following conditions on a real $n \times n$ matrix A are equivalent:

(a) A is orthogonal.
(b) Multiplication by A preserves dot product, that is, $(AX \cdot AY) = (X \cdot Y)$ for all column vectors X, Y.
(c) The columns of A are mutually orthogonal unit vectors.

A basis consisting of mutually orthogonal unit vectors is called an *orthonormal* basis. An orthogonal matrix is one whose columns form an orthonormal basis.

Left multiplication by an orthogonal matrix is also called an *orthogonal operator*. Thus the orthogonal operators on $\mathbb{R}^n$ are the ones which preserve dot product.

Proof of Proposition (5.13). We write $(X \cdot Y) = X^t Y$. If A is orthogonal, then $A^t A = I$, so

$$(X \cdot Y) = X^t Y = X^t A^t A Y = (AX)^t (AY) = (AX \cdot AY).$$

Conversely, suppose that $X^t Y = X^t A^t A Y$ for all X and Y. We rewrite this equality as $X^t B Y = 0$, where $B = I - A^t A$. For any matrix B,

(5.14) $$e_i{}^t B e_j = b_{ij}.$$

So if $X^t B Y = 0$ for all X, Y, then $e_i{}^t B e_j = b_{ij} = 0$ for all i, j, and $B = 0$. Therefore $I = A^t A$. This proves the equivalence of (a) and (b). To prove that (a) and (c) are equivalent, let A_j denote the jth column of the matrix A. The (i, j) entry of the product matrix $A^t A$ is $(A_i \cdot A_j)$. Thus $A^t A = I$ if and only if $(A_i \cdot A_i) = 1$ for all i,

and $(A_i \cdot A_j) = 0$ for all $i \neq j$, which is to say that the columns have length 1 and are orthogonal. □

The geometric meaning of multiplication by an orthogonal matrix can be explained in terms of rigid motions. A *rigid motion* or *isometry* of $\mathbb{R}^n$ is a map $m: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ which is distance preserving; that is, it is a map satisfying the following condition: If X, Y are points of $\mathbb{R}^n$, then the distance from X to Y is equal to the distance from $m(X)$ to $m(Y)$:

$$(5.15) \qquad |m(X) - m(Y)| = |X - Y|.$$

Such a rigid motion carries a triangle to a congruent triangle, and therefore it preserves angles and shapes in general.

Note that the composition of two rigid motions is a rigid motion, and that the inverse of a rigid motion is a rigid motion. Therefore the rigid motions of $\mathbb{R}^n$ form a group M_n, with composition of operations as its law of composition. This group is called the *group of motions*.

(5.16) **Proposition.** Let m be a map $\mathbb{R}^n \longrightarrow \mathbb{R}^n$. The following conditions on m are equivalent:

(a) m is a rigid motion which fixes the origin.
(b) m preserves dot product; that is, for all $X, Y \in \mathbb{R}^n$, $(m(X) \cdot m(Y)) = (X \cdot Y)$.
(c) m is left multiplication by an orthogonal matrix.

(5.17) **Corollary.** A rigid motion which fixes the origin is a linear operator.

This follows from the equivalence of (a) and (c).

Proof of Proposition (5.16). We will use the shorthand $'$ to denote the map m, writing $m(X) = X'$. Suppose that m is a rigid motion fixing 0. With the shorthand notation, the statement (5.15) that m preserves distance reads

$$(5.18) \qquad (X' - Y' \cdot X' - Y') = (X - Y \cdot X - Y)$$

for all vectors X, Y. Setting $Y = 0$ shows that $(X' \cdot X') = (X \cdot X)$ for all X. We expand both sides of (5.18) and cancel $(X \cdot X)$ and $(Y \cdot Y)$, obtaining $(X' \cdot Y') = (X \cdot Y)$. This shows that m preserves dot product, hence that (a) implies (b).

To prove that (b) implies (c), we note that the only map which preserves dot product and which also fixes each of the basis vectors e_i is the identity. For, if m preserves dot product, then $(X \cdot e_j) = (X' \cdot e_j')$ for any X. If $e_j' = e_j$ as well, then

$$x_j = (X \cdot e_j) = (X' \cdot e_j') = (X' \cdot e_j) = x_j'$$

for all j. Hence $X = X'$, and m is the identity.

Now suppose that m preserves dot product. Then the images $e_1', \dots, e_n'$ of the standard basis vectors are orthonormal: $(e_i' \cdot e_i') = 1$ and $(e_i' \cdot e_j') = 0$ if $i \neq j$. Let $\mathbf{B}' = (e_1', \dots, e_n')$, and let $A = [\mathbf{B}']$. According to Proposition (5.13), A is an or-

thogonal matrix. Since the orthogonal matrices form a group, A^{-1} is also orthogonal. This being so, multiplication by A^{-1} preserves dot product too. So the composed motion $A^{-1}m$ preserves dot product, and it fixes each of the basis vectors e_i. Therefore $A^{-1}m$ is the identity map. This shows that m is left multiplication by A, as required.

Finally, if m is a linear operator whose matrix A is orthogonal, then $X' - Y' = (X - Y)'$ because m is linear, and $|X' - Y'| = |(X - Y)'| = |X - Y|$ by (5.13b). So m is a rigid motion. Since a linear operator also fixes 0, this shows that (c) implies (a). □

One class of rigid motions which do not fix the origin, and which are therefore not linear operators, is the translations. Given any fixed vector $b = (b_1,\ldots,b_n)^t$ in $\mathbb{R}^n$, *translation* by b is the map

(5.19)
$$t_b(X) = X + b = \begin{bmatrix} x_1 + b_1 \\ \vdots \\ x_n + b_n \end{bmatrix}.$$

This map is a rigid motion because $t_b(X) - t_b(Y) = (X + b) - (Y + b) = X - Y$, and hence $|t_b(X) - t_b(Y)| = |X - Y|$.

(5.20) **Proposition.** Every rigid motion m is the composition of an orthogonal linear operator and a translation. In other words, it has the form $m(X) = AX + b$ for some orthogonal matrix A and some vector b.

Proof. Let $b = m(0)$. Then $t_{-b}(b) = 0$, so the composed operation $t_{-b}m$ is a rigid motion which fixes the origin: $t_{-b}(m(0)) = 0$. According to Proposition (5.16), $t_{-b}m$ is left multiplication by an orthogonal matrix A: $t_{-b}m(X) = AX$. Applying t_b to both sides of this equation, we find $m(X) = AX + b$.

Note that both the vector b and the matrix A are uniquely determined by m, because $b = m(0)$ and A is the operator $t_{-b}m$. □

Recall that the determinant of an orthogonal matrix is ± 1. An orthogonal operator is called *orientation-preserving* if its determinant is $+1$, and *orientation-reversing* if its determinant is -1. Similarly, let m be a rigid motion. We write $m(X) = AX + b$ as above. Then m is called *orientation-preserving* if $\det A = 1$, and *orientation-reversing* if $\det A = -1$. A motion of $\mathbb{R}^2$ is orientation-reversing if it flips the plane over, and orientation-preserving if it does not.

Combining Theorem (5.5) with Proposition (5.16) gives us the following characterization of rotations:

(5.21) **Corollary.** The rotations of $\mathbb{R}^2$ and $\mathbb{R}^3$ are the orientation-preserving rigid motions which fix the origin. □

We now proceed to the proof of Theorem (5.5), which characterizes the rotations of $\mathbb{R}^2$ and $\mathbb{R}^3$ about the origin. Every rotation ρ is a rigid motion, so Proposi-

tion (5.16) tells us that ρ is multiplication by an orthogonal matrix A. Also, the determinant of A is 1. This is because $\det A = \pm 1$ for any orthogonal matrix, and because the determinant varies continuously with the angle of rotation. When the angle is zero, A is the identity matrix, which has determinant 1. Thus the matrix of a rotation is an element of SO_2 or SO_3.

Conversely, let $A \in SO_2$ be an orthogonal 2×2 matrix of determinant 1. Let v_1 denote the first column Ae_1 of A. Since A is orthogonal, v_1 is a unit vector. There is a rotation R (3.1) such that $Re_1 = v_1$ too. Then $B = R^{-1}A$ fixes e_1. Also, A and R are elements of SO_2, and this implies that B is in SO_2. So the columns of B form an orthonormal basis of $\mathbb{R}^2$, and the first column is e_1. Being of length 1 and orthogonal to e_1, the second column must be either e_2 or $-e_2$, and the second case is ruled out by the fact that $\det B = 1$. It follows that $B = I$ and that $A = R$. So A is a rotation.

To prove that an element A of SO_3 represents a rotation, we'd better decide on a definition of a rotation ρ of $\mathbb{R}^3$ about the origin. We will require the following:

(5.22)

(i) ρ is a rigid motion which fixes the origin;
(ii) ρ also fixes a nonzero vector v;
(iii) ρ operates as a rotation on the plane P orthogonal to v.

According to Proposition (5.16), the first condition is equivalent to saying that ρ is an orthogonal operator. So our matrix $A \in SO_3$ satisfies this condition. Condition (ii) can be stated by saying that v is an eigenvector for the operator ρ, with eigenvalue 1. Then since ρ preserves orthogonality, it sends the orthogonal space P to itself. In other words, P is an invariant subspace. Condition (iii) says that the restriction of ρ to this invariant subspace is a rotation.

Notice that the matrix (5.2) does satisfy these conditions, with $v = e_1$.

(5.23) **Lemma.** Every element $A \in SO_3$ has the eigenvalue 1.

Proof. We will show that $\det(A - I) = 0$. This will prove the lemma [see (4.8)]. This proof is tricky, but efficient. Recall that $\det A = \det A^t$ for any matrix A, so $\det A^t = 1$. Since A is orthogonal, $A^t(A - I) = (I - A)^t$. Then

$$\det(A - I) = \det A^t(A - I) = \det(I - A)^t = \det(I - A).$$

On the other hand, for any 3×3 matrix B, $\det(-B) = -\det B$. Therefore $\det(A - I) = -\det(I - A)$, and it follows that $\det(A - I) = 0$. □

Now given a matrix $A \in SO_3$, the lemma shows that left multiplication by A fixes a nonzero vector v_1. We normalize its length to 1, and we choose orthogonal unit vectors v_2, v_3 lying in the plane P orthogonal to v_1. Then $\mathbf{B} = (v_1, v_2, v_3)$ is an orthonormal basis of $\mathbb{R}^3$. The matrix $P = [\mathbf{B}]^{-1}$ is orthogonal because $[\mathbf{B}]$ is orthogo-

nal, and $A' = PAP^{-1}$ represents the same operator as A does, with respect to the basis $\mathbf{B}$. Since A and P are orthogonal, so is A'. Also $\det A' = \det A = 1$. So $A' \in SO_3$.

Since v_1 is an eigenvector with eigenvalue 1, the first column of A' is e_1. Since A' is orthogonal, the other columns are orthogonal to e_1, and A' has the block form

$$\left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & R \end{array}\right].$$

Using the fact that $A' \in SO_3$, one finds that $R \in SO_2$. So R is a rotation. This shows that A' has the form (5.2) and that it represents a rotation. Hence A does too. This completes the proof of Theorem (5.5). □

(5.24) *Note.* To keep the new basis separate from the old basis, we denoted it by $\mathbf{B}'$ in Chapter 3. The prime is not needed when the old basis is the standard basis, and since it clutters the notation, we will often drop it, as we did here.

6. DIAGONALIZATION

In this section we show that for "most" linear operators on a *complex* vector space, there is a basis such that the matrix of the operator is diagonal. The key fact, which we already noted at the end of Section 4, is that every complex polynomial of positive degree has a root. This tells us that every linear operator has an eigenvector.

(6.1) **Proposition.**

(a) *Vector space form:* Let T be a linear operator on a finite-dimensional complex vector space V. There is a basis $\mathbf{B}$ of V such that the matrix A of T is upper triangular.

(b) *Matrix form:* Every complex $n \times n$ matrix A is similar to an upper triangular matrix. In other words, there is a matrix $P \in GL_n(\mathbb{C})$ such that PAP^{-1} is upper triangular.

Proof. The two assertions are equivalent, because of (3.5). We begin by applying (4.19b), which shows the existence of an eigenvector, call it v_1'. Extend to a basis $\mathbf{B}' = (v_1', \dots, v_n')$ for V. Then by (3.11), the first column of the matrix A' of T with respect to $\mathbf{B}'$ will be $(c_1, 0, \dots, 0)^t$, where c_1 is the eigenvalue of v_1'. Therefore A' has the form

$$A' = \left[\begin{array}{c|ccc} c_1 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right]$$

where B is an $(n-1) \times (n-1)$ matrix. The matrix version of this reduction is this: Given any $n \times n$ matrix A, there is a $P \in GL_n(\mathbb{C})$ such that $A' = PAP^{-1}$ has the above form. Now apply induction on n. By induction, we may assume that the existence of some $Q \in GL_{n-1}(\mathbb{C})$ such that QBQ^{-1} is triangular has been proved. Let Q_1 be the $n \times n$ matrix

$$\left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{array}\right].$$

Then

$$(Q_1P)A(Q_1P)^{-1} = Q_1(PAP^{-1})Q_1^{-1} = Q_1A'Q_1^{-1}$$

has the form

$$\left[\begin{array}{c|ccc} c_1 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & QBQ^{-1} & \\ 0 & & & \end{array}\right],$$

which is triangular. □

As we mentioned, the important point in the proof is that every complex polynomial has a root. The same proof will work for any field F, provided that all the roots of the characteristic polynomial are in the field.

(6.2) **Corollary.** Let F be a field.

(a) *Vector space form:* Let T be a linear operator on a finite-dimensional vector space V over F, and suppose that the characteristic polynomial of T factors into linear factors in the field F. Then there is a basis **B** of V such that the matrix A of T is triangular.

(b) *Matrix form:* Let A be an $n \times n$ matrix whose characteristic polynomial factors into linear factors in the field F. There is a matrix $P \in GL_n(F)$ such that PAP^{-1} is triangular.

Proof. The proof is the same, except that to make the induction step one has to check that the characteristic polynomial of the matrix B is $p(t)/(t - c_1)$, where $p(t)$ is the characteristic polynomial of A. This is true because $p(t)$ is also the characteristic polynomial of A' (4.17), and because $\det(tI - A') = (t - c_1)\det(tI - B)$.

So our hypothesis that the characteristic polynomial factors into linear factors carries over from A to B. □

Let us now ask which matrices A are similar to *diagonal* matrices. As we saw in (3.12), these are the matrices A which have a basis of eigenvectors. Suppose again that $F = \mathbb{C}$, and look at the roots of the characteristic polynomial $p(t)$. Each root is the eigenvalue associated to some eigenvector, and an eigenvector has only one eigenvalue. Most complex polynomials of degree n have n distinct roots. So most complex matrices have n eigenvectors with different eigenvalues, and it is reasonable to suppose that these eigenvectors may form a basis. This is true.

(6.3) **Proposition.** Let $v_1,\dots,v_r \in V$ be eigenvectors for a linear operator T, with distinct eigenvalues $c_1,\dots,c_r$. Then the set $(v_1,\dots,v_r)$ is linearly independent.

Proof. Induction on r: Suppose that a dependence relation

$$0 = a_1v_1 + \cdots + a_rv_r$$

is given. We must show that $a_i = 0$ for all i, and to do so we apply the operator T:

$$0 = T(0) = a_1T(v_1) + \cdots + a_rT(v_r) = a_1c_1v_1 + \cdots + a_rc_rv_r.$$

This is a second dependence relation among $(v_1,\dots,v_r)$. We eliminate v_r from the two relations, multiplying the first relation by c_r and subtracting the second:

$$0 = a_1(c_r - c_1)v_1 + \cdots + a_{r-1}(c_r - c_{r-1})v_{r-1}.$$

Applying the principle of induction, we assume that $(v_1,\dots,v_{r-1})$ are independent. Then the coefficients $a_1(c_r - c_1),\dots,a_{r-1}(c_r - c_{r-1})$ are all zero. Since the c_i's are distinct, $c_r - c_i \neq 0$ if $i < r$. Thus $a_1 = \dots = a_{r-1} = 0$, and the original relation is reduced to $0 = a_rv_r$. Since an eigenvector can not be zero, $a_r = 0$ too. □

The next theorem follows by combining (3.12) and (6.3):

(6.4) **Theorem.** Let T be a linear operator on a vector space V of dimension n over a field F. Assume that its characteristic polynomial has n distinct roots in F. Then there is a basis for V with respect to which the matrix of T is diagonal. □

Note that the diagonal entries are determined, except for their order, by the linear operator T. They are the eigenvalues.

When $p(t)$ has multiple roots, there is usually no basis of eigenvectors, and it is harder to find a nice matrix for T. The study of this case leads to what is called the *Jordan canonical form* for a matrix, which will be discussed in Chapter 12.

As an example of diagonalization, consider the matrix

$$A = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$$

whose eigenvectors were computed in (4.10). These eigenvectors form a basis $\mathbf{B} = (v_1, v_2)$ of $\mathbb{R}^2$. According to [Chapter 3 (4.20), see also Note (5.24)], the matrix relating the standard basis $\mathbf{E}$ to this basis $\mathbf{B}$ is

(6.5)
$$P = [\mathbf{B}]^{-1} = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}^{-1} = -\frac{1}{3}\begin{bmatrix} -1 & -2 \\ -1 & 1 \end{bmatrix},$$

and $PAP^{-1} = A'$ is diagonal:

(6.6)
$$-\frac{1}{3}\begin{bmatrix} -1 & -2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 5 & \\ & 2 \end{bmatrix} = A'.$$

The general rule is stated in Corollary (6.7):

(6.7) **Corollary.** If a basis $\mathbf{B}$ of eigenvectors of A in F^n is known and if $P = [\mathbf{B}]^{-1}$, then $A' = PAP^{-1}$ is diagonal. □

The importance of Theorem (6.4) comes from the fact that it is easy to compute with diagonal matrices. For example, if $A' = PAP^{-1}$ is diagonal, then we can compute powers of the matrix A using the formula

(6.8)
$$A^k = (P^{-1}A'P)^k = P^{-1}A'^k P.$$

Thus if A is the matrix (4.9), then

$$A^k = -\frac{1}{3}\begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 5 & \\ & 2 \end{bmatrix}^k\begin{bmatrix} -1 & -2 \\ -1 & 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 5^k + 2 \cdot 2^k & 2(5^k - 2^k) \\ 5^k - 2^k & 2 \cdot 5^k + 2^k \end{bmatrix}.$$

7. *SYSTEMS OF DIFFERENTIAL EQUATIONS*

We learn in calculus that the solutions to the first-order linear differential equation

(7.1)
$$\frac{dx}{dt} = ax$$

are $x(t) = ce^{at}$, c being an arbitrary constant. Indeed, ce^{at} obviously solves (7.1). To show that every solution has this form, let $x(t)$ be an arbitrary differentiable function which is a solution. We differentiate $e^{-at}x(t)$ using the product rule:

$$\frac{d}{dt}(e^{-at}x(t)) = -ae^{-at}x(t) + e^{-at}ax(t) = 0.$$

Thus $e^{-at}x(t)$ is a constant c, and $x(t) = ce^{at}$.

As an application of diagonalization, we will extend this solution to systems of differential equations. In order to write our equations in matrix notation, we use the following terminology. A *vector-valued function* $X(t)$ is a vector whose entries are

functions of t. Similarly, a *matrix-valued function* $A(t)$ is a matrix whose entries are functions:

$$X(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \quad A(t) = \begin{bmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \vdots & & \vdots \\ a_{m1}(t) & \cdots & a_{mn}(t) \end{bmatrix}. \tag{7.2}$$

The calculus operations of taking limits, differentiating, and so on are extended to vector-valued and matrix-valued functions by performing the operations on each entry separately. Thus by definition

$$\lim_{t \to t_0} X(t) = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}, \quad \text{where } \xi_i = \lim_{t \to t_0} x_i(t). \tag{7.3}$$

So this limit exists if and only if $\lim x_i(t)$ exists for each i. Similarly, the derivative of a vector-valued or matrix-valued function is the function obtained by differentiating each entry separately:

$$\frac{dX}{dt} = \begin{bmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{bmatrix}, \quad \frac{dA}{dt} = \begin{bmatrix} a_{11}'(t) & \cdots & a_{1n}'(t) \\ \vdots & & \vdots \\ a_{m1}'(t) & \cdots & a_{mn}'(t) \end{bmatrix},$$

where $x_i'(t)$ is the derivative of $x_i(t)$, and so on. So dX/dt is defined if and only if each of the functions $x_i(t)$ is differentiable. The derivative can also be described in vector notation, as

$$\frac{dX}{dt} = \lim_{h \to 0} \frac{X(t+h) - X(t)}{h}. \tag{7.4}$$

Here $X(t+h) - X(t)$ is computed by vector addition and the h in the denominator stands for scalar multiplication by h^{-1}. The limit is obtained by evaluating the limit of each entry separately, as above. So the entries of (7.4) are the derivatives $x_i'(t)$. The same is true for matrix-valued functions.

A system of homogeneous first-order linear, constant-coefficient differential equations is a matrix equation of the form

$$\frac{dX}{dt} = AX, \tag{7.5}$$

where A is an $n \times n$ real or complex matrix and $X(t)$ is an n-dimensional vector-valued function. Writing out such a system, we obtain a system of n differential

equations, of the form

$$\begin{aligned} \frac{dx_1}{dt} &= a_{11}x_1(t) + \cdots + a_{1n}x_n(t) \\ &\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ \frac{dx_n}{dt} &= a_{n1}x_1(t) + \cdots + a_{nn}x_n(t). \end{aligned} \tag{7.6}$$

The $x_i(t)$ are unknown functions, and the a_{ij} are scalars. For example, if we substitute the matrix $\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$ for A, (7.5) becomes a system of two equations in two unknowns:

$$\begin{aligned} \frac{dx_1}{dt} &= 3x_1 + 2x_2 \\ \frac{dx_2}{dt} &= x_1 + 4x_2. \end{aligned} \tag{7.7}$$

The simplest systems (7.5) are those in which A is a diagonal matrix. Let the diagonal entries be a_i. Then equation (7.6) reads

$$\frac{dx_i}{dt} = a_i x_i(t), \quad i = 1,\ldots,n. \tag{7.8}$$

Here the unknown functions x_i are not mixed up by the equations, so we can solve for each one separately:

$$x_i = c_i e^{a_i t}, \tag{7.9}$$

for some constant c_i.

The observation which allows us to solve the differential equation (7.5) in most cases is this: If v is an eigenvector for A with eigenvalue a, then

$$X = e^{at}v \tag{7.10}$$

is a particular solution of (7.5). Here $e^{at}v$ is to be interpreted as the scalar product of the function e^{at} and the vector v. Differentiation operates on the scalar function, fixing the constant vector v, while multiplication by A operates on the vector v, fixing the scalar function e^{at}. Thus $\frac{d}{dt}e^{at}v = ae^{at}v = Ae^{at}v$. For example, $(2,-1)^t$ is an eigenvector with eigenvalue 2 of the matrix $\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$, and $\begin{bmatrix} 2e^{2t} \\ -e^{2t} \end{bmatrix}$ solves the system of differential equations (7.7).

This observation allows us to solve (7.5) whenever the matrix A has distinct real eigenvalues. In that case every solution will be a linear combination of the special solutions (7.10). To work this out, it is convenient to diagonalize. Let us replace

the notation $'$ used in the previous section by $\tilde{}$ here, to avoid confusion with differentiation. Let P be an invertible matrix such that $PAP^{-1} = \tilde{A}$ is diagonal. So $P = [\mathbf{B}]^{-1}$, where $\mathbf{B}$ is a basis of eigenvectors. We make the linear change of variable

(7.11) $$X = P^{-1}\tilde{X}.$$

Then

(7.12) $$\frac{dX}{dt} = P^{-1}\frac{d\tilde{X}}{dt}.$$

Substituting into (7.5), we find

(7.13) $$\frac{d\tilde{X}}{dt} = PAP^{-1}\tilde{X} = \tilde{A}\tilde{X}.$$

Since $\tilde{A}$ is diagonal, the variables $\tilde{x}_i$ have been separated, so the equation can be solved in terms of exponentials. The diagonal entries of $\tilde{A}$ are the eigenvalues $\lambda_1,\dots,\lambda_n$ of A, so the solution of the system (7.13) is

(7.14) $$\tilde{x}_i = c_i e^{\lambda_i t}, \quad \text{for some } c_i.$$

Substituting back,

(7.15) $$X = P^{-1}\tilde{X}$$

solves the original system (7.5). This proves the following:

(7.16) **Proposition.** Let A be an $n \times n$ matrix, and let P be an invertible matrix such that $PAP^{-1} = \tilde{A}$ is diagonal, with diagonal entries $\lambda_1,\dots,\lambda_n$. The general solution of the system $\frac{dX}{dt} = AX$ is $X = P^{-1}\tilde{X}$, where $\tilde{x}_i = c_i e^{\lambda_i t}$, for some arbitrary constants c_i. □

The matrix which diagonalizes A in example (7.7) was computed in (6.5):

(7.17) $$P^{-1} = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}, \quad \text{and} \quad \tilde{A} = \begin{bmatrix} 5 & \\ & 2 \end{bmatrix}.$$

Thus

$$\begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{bmatrix} = \begin{bmatrix} c_1 e^{5t} \\ c_2 e^{2t} \end{bmatrix} \quad \text{and}$$

(7.18) $$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} c_1 e^{5t} \\ c_2 e^{2t} \end{bmatrix} = \begin{bmatrix} c_1 e^{5t} + 2c_2 e^{2t} \\ c_1 e^{5t} - c_2 e^{2t} \end{bmatrix}.$$

In other words, every solution is a linear combination of the two basic solutions

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} e^{5t} \\ e^{5t} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2e^{2t} \\ -e^{2t} \end{bmatrix}.$$

These are the solutions (7.10) corresponding to the eigenvectors $(1, 1)^t$ and $(2, -1)^t$. The coefficients c_i appearing in these solutions are arbitrary. They are usually determined by assigning *initial conditions*, meaning the value of X at some particular t_0.

Let us now consider the case that the coefficient matrix A has distinct eigenvalues, but that they are not all real. To copy the method which we used above, we must first consider differential equations of the form (7.1), in which a is a complex number. Properly interpreted, the solutions of such a differential equation still have the form ce^{at}. The only thing to remember is that e^{at} will now be a complex-valued function of t. In order to focus attention, we restrict the variable t to real values here, although this is not the most natural choice when working with complex-valued functions. Allowing t to take on complex values would not change things very much.

The definition of the derivative of a complex-valued function is the same as for real-valued functions:

$$\frac{dx}{dt} = \lim_{h \to 0} \frac{x(t+h) - x(t)}{h}, \tag{7.19}$$

provided that this limit exists. There are no new features. We can write any such function $x(t)$ in terms of its real and imaginary parts, which will be real-valued functions:

$$x(t) = u(t) + iv(t). \tag{7.20}$$

Then x is differentiable if and only if u and v are differentiable, and if they are, the derivative of x is $x' = u' + iv'$. This follows directly from the definition. The usual rules for differentiation, such as the product rule, hold for complex-valued functions. These rules can be proved by applying the corresponding theorem for real functions to u and v, or else by carrying the proof for real functions over to the complex case.

Recall the formula

$$e^{r+si} = e^r(\cos s + i \sin s). \tag{7.21}$$

Differentiation of this formula shows that $de^{at}/dt = ae^{at}$ for all complex numbers $a = r + si$. Therefore ce^{at} solves the differential equation (7.1), and the proof given at the beginning of the section shows that these are the only solutions.

Having extended the case of one equation to complex coefficients, we can now use the method of diagonalization to solve a system of equations (7.5) when A is an arbitrary complex matrix with distinct eigenvalues.

For example, let $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$. The vectors $v_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}$ and $v_2 = \begin{bmatrix} i \\ 1 \end{bmatrix}$ are eigenvectors, with eigenvalues $1 + i$ and $1 - i$ respectively. Let $\mathbf{B} = (v_1, v_2)$. According to (6.7), A is diagonalized by the matrix P, where

$$P^{-1} = [\mathbf{B}] = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}. \tag{7.22}$$

Formula (7.14) tells us that $\tilde{X} = \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{bmatrix} = \begin{bmatrix} c_1 e^{t+it} \\ c_2 e^{t-it} \end{bmatrix}$. The solutions of (7.5) are

(7.23) $$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = P^{-1}\tilde{X} = \begin{bmatrix} c_1 e^{t+it} + ic_2 e^{t-it} \\ ic_1 e^{t+it} + c_2 e^{t-it} \end{bmatrix},$$

where c_1, c_2 are arbitrary complex numbers. So every solution is a linear combination of the two basic solutions

(7.24) $$\begin{bmatrix} e^{t+it} \\ ie^{t+it} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} ie^{t-it} \\ e^{t-it} \end{bmatrix}.$$

However, these solutions are not completely satisfactory, because we began with a system of differential equations with real coefficients, and the answer we obtained is complex. When the original matrix is real, we want to have real solutions. We note the following lemma:

(7.25) **Lemma.** Let A be a real $n \times n$ matrix, and let $X(t)$ be a complex-valued solution of the differential equation (7.5). The real and imaginary parts of $X(t)$ solve the same equation. □

Now *every* solution of the original equation (7.5), whether real or complex, has the form (7.23) for some complex numbers c_i. So the real solutions are among those we have found. To write them down explicitly, we may take the real and imaginary parts of the complex solutions.

The real and imaginary parts of the basic solutions (7.24) are determined using (7.21). They are

(7.26) $$\begin{bmatrix} e^t \cos t \\ -e^t \sin t \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} e^t \sin t \\ e^t \cos t \end{bmatrix}.$$

Every real solution is a real linear combination of these particular solutions.

8. THE MATRIX EXPONENTIAL

Systems of first-order linear, constant-coefficient differential equations can also be solved formally, using the *matrix exponential*. The exponential of an $n \times n$ real or complex matrix A is obtained by substituting a matrix into the Taylor's series

(8.1) $$1 + x/1! + x^2/2! + x^3/3! + \cdots$$

for e^x. Thus by definition,

(8.2) $$e^A = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots.$$

This is an $n \times n$ matrix.

(8.3) **Proposition.** The series (8.2) converges absolutely for all complex matrices A.

In order not to break up the discussion, we have collected the proofs together at the end of the section.

Since matrix multiplication is relatively complicated, it isn't easy to write down the matrix entries of e^A directly. In particular, the entries of e^A are usually not obtained by exponentiating the entries of A. But one case in which they are, and in which the exponential is easily computed, is when A is a diagonal matrix, say with diagonal entries a_i. Inspection of the series shows that e^A is also diagonal in this case and that its diagonal entries are e^{a_i}.

The exponential is also relatively easy to compute for a triangular 2×2 matrix. For example, let

$$(8.4) \qquad A = \begin{bmatrix} 1 & 1 \\ & 2 \end{bmatrix}.$$

Then

$$(8.5) \qquad e^A = \begin{bmatrix} 1 & \\ & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ & 2 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 1 & 3 \\ & 4 \end{bmatrix} + \cdots = \begin{bmatrix} e & * \\ & e^2 \end{bmatrix}.$$

The diagonal entries are exponentiated to obtain the diagonal entries of e^A. It is a good exercise to calculate the missing entry $*$ directly from the definition.

The exponential of a matrix A can also be determined whenever we know a matrix P such that PAP^{-1} is diagonal. Using the rule $PA^kP^{-1} = (PAP^{-1})^k$ and the distributive law for matrix multiplication, we find

$$(8.6) \qquad Pe^AP^{-1} = PIP^{-1} + (PAP^{-1}) + \frac{1}{2!}(PAP^{-1})^2 + \cdots = e^{PAP^{-1}}.$$

Suppose that $PAP^{-1} = \tilde{A}$ is diagonal, with diagonal entries λ_i. Then $e^{\tilde{A}}$ is also diagonal, and its diagonal entries are e^{λ_i}. Therefore we can compute e^A explicitly:

$$(8.7) \qquad e^A = P^{-1}e^{\tilde{A}}P.$$

In order to use the matrix exponential to solve systems of differential equations, we need to extend some of the properties of the ordinary exponential to it. The most fundamental property is $e^{x+y} = e^xe^y$. This property can be expressed as a formal identity between the two infinite series which are obtained by expanding

$$(8.8) \qquad \begin{aligned} e^{x+y} &= 1 + (x+y)/1! + (x+y)^2/2! + \cdots \quad \text{and} \\ e^xe^y &= (1 + x/1! + x^2/2! + \cdots)(1 + y/1! + y^2/2! + \cdots). \end{aligned}$$

We can not substitute matrices into this identity because the commutative law is needed to obtain equality of the two series. For instance, the quadratic terms of (8.8), computed without the commutative law, are $\frac{1}{2}(x^2 + xy + yx + y^2)$ and $\frac{1}{2}x^2 + xy + \frac{1}{2}y^2$. They are not equal unless $xy = yx$. So there is no reason to expect

e^{A+B} to equal $e^A e^B$ in general. However, if two matrices A and B happen to commute, the formal identity can be applied.

(8.9) **Proposition.**

(a) The formal expansions of (8.8), with commuting variables x,y, are equal.
(b) Let A,B be complex $n \times n$ matrices which commute: $AB = BA$. Then $e^{A+B} = e^A e^B$.

The proof is at the end of the section. □

(8.10) **Corollary.** For any $n \times n$ complex matrix A, the exponential e^A is invertible, and its inverse is e^{-A}.

This follows from the proposition because A and $-A$ commute, and hence $e^A e^{-A} = e^{A-A} = e^0 = I$. □

As a sample application of Proposition (8.9b), consider the matrix

$$A = \begin{bmatrix} 2 & 3 \\ & 2 \end{bmatrix}. \tag{8.11}$$

We can compute its exponential by writing it in the form $A = 2I + B$, where $B = 3e_{12}$. Since $2I$ commutes with B, Proposition (8.9b) applies: $e^A = e^{2I}e^B$, and from the series expansion we read off the values $e^{2I} = e^2 I$ and $e^B = I + B$. Thus

$$e^A = \begin{bmatrix} e^2 & \\ & e^2 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ & 1 \end{bmatrix} = \begin{bmatrix} e^2 & 3e^2 \\ & e^2 \end{bmatrix}.$$

We now come to the main result relating the matrix exponential to differential equations. Given an $n \times n$ matrix A, we consider the exponential e^{tA}, t being a variable scalar, as a matrix-valued function:

$$e^{tA} = I + tA + \frac{t^2}{2!}A^2 + \frac{t^3}{3!}A^3 + \cdots. \tag{8.12}$$

(8.13) **Proposition.** e^{tA} is a differentiable function of t, and its derivative is Ae^{tA}.

The proof is at the end of the section. □

(8.14) **Theorem.** Let A be a real or complex $n \times n$ matrix. The columns of the matrix e^{tA} form a basis for the vector space of solutions of the differential equation

$$\frac{dX}{dt} = AX.$$

We will need the following lemma, whose proof is an exercise:

(8.15) **Lemma.** *Product rule:* Let $A(t)$ and $B(t)$ be differentiable matrix-valued functions of t, of suitable sizes so that their product is defined. Then the matrix product $A(t)B(t)$ is differentiable, and its derivative is

$$\frac{d}{dt}(A(t)B(t)) = \frac{dA}{dt}B + A\frac{dB}{dt}. \square$$

Proof of Theorem (8.14). Proposition (8.13) shows that the columns of A solve the differential equation, because differentiation and multiplication by A act independently on the columns of the matrix e^{tA}. To show that every solution is a linear combination of the columns, we copy the proof given at the beginning of Section 7. Let $X(t)$ be an arbitrary solution of (7.5). We differentiate the matrix product $e^{-tA}X(t)$, obtaining

$$\frac{d}{dt}(e^{-tA}X(t)) = -Ae^{-tA}X(t) + e^{-tA}AX(t).$$

Fortunately, A and e^{-tA} commute. This follows directly from the definition of the exponential. So the derivative is zero. Therefore, $e^{-tA}X(t)$ is a constant column vector, say $C = (c_1,\ldots,c_n)^t$, and $X(t) = e^{tA}C$. This expresses $X(t)$ as a linear combination of the columns of e^{tA}. The expression is unique because e^{tA} is an invertible matrix. $\square$

According to Theorem (8.14), the matrix exponential always solves the differential equation (7.5). Since direct computation of the exponential can be quite difficult, this theorem may not be easy to apply in a concrete situation. But if A is a diagonalizable matrix, then the exponential can be computed as in (8.7): $e^A = P^{-1}e^{\tilde{A}}P$. We can use this method of evaluating e^{tA} to solve equation (7.5), but of course it gives the same result as before. Thus if A is the matrix used in example (7.7), so that $P, \tilde{A}$ are as in (7.17), then

$$e^{t\tilde{A}} = \begin{bmatrix} e^{5t} & \\ & e^{2t} \end{bmatrix}$$

and

$$e^{tA} = P^{-1}e^{t\tilde{A}}P = -\frac{1}{3}\begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} e^{5t} & \\ & e^{2t} \end{bmatrix}\begin{bmatrix} -1 & -2 \\ -1 & 1 \end{bmatrix}$$

$$= \frac{1}{3}\begin{bmatrix} e^{5t} + 2e^{2t} & 2e^{5t} - 2e^{2t} \\ e^{5t} - e^{2t} & 2e^{5t} + e^{2t} \end{bmatrix}.$$

The columns we have obtained form a second basis for the general solution (7.18).

On the other hand, the matrix $A = \begin{bmatrix} 1 & \\ 1 & 1 \end{bmatrix}$, which represents the system of equations

(8.16) $$\frac{dx}{dt} = x, \quad \frac{dy}{dt} = x + y,$$

is not diagonalizable. So the method of Section 7 can not be applied. To solve it, we write $At = It + Bt$, where $B = e_{21}$, and find, as in the discussion of (8.11),

$$e^{At} = e^{It}e^{Bt} = \begin{bmatrix} e^t & \\ te^t & e^t \end{bmatrix}. \tag{8.17}$$

Thus the solutions of (8.16) are linear combinations of the columns

$$\begin{bmatrix} e^t \\ te^t \end{bmatrix}, \begin{bmatrix} 0 \\ e^t \end{bmatrix}. \tag{8.18}$$

To compute the exponential explicitly in all cases requires putting the matrix into Jordan form (see Chapter 12).

We now go back to prove Propositions (8.3), (8.9), and (8.13). For want of a more compact notation, we will denote the i, j-entry of a matrix A by A_{ij} here. So $(AB)_{ij}$ will stand for the entry of the product matrix AB, and $(A^k)_{ij}$ for the entry of A^k. With this notation, the i, j-entry of e^A is the sum of the series

$$(e^A)_{ij} = I_{ij} + A_{ij} + \frac{1}{2!}(A^2)_{ij} + \frac{1}{3!}(A^3)_{ij} + \cdots. \tag{8.19}$$

In order to prove that the series for the exponential converges, we need to show that the entries of the powers A^k of a given matrix do not grow too fast, so that the absolute values of the i, j-entries form a bounded (and hence convergent) series. Let us define the *norm* of an $n \times n$ matrix A to be the maximum absolute value of the matrix entries:

$$\|A\| = \max_{i,j} |A_{ij}|. \tag{8.20}$$

In other words, $\|A\|$ is the smallest real number such that

$$|A_{ij}| \leq \|A\| \quad \text{for all } i,j. \tag{8.21}$$

This is one of several possible definitions of the norm. Its basic property is as follows:

(8.22) **Lemma.** Let A, B be complex $n \times n$ matrices. Then $\|AB\| \leq n\|A\|\|B\|$, and $\|A^k\| \leq n^{k-1}\|A\|^k$ for all $k > 0$.

Proof. We estimate the size of the i, j-entry of AB:

$$|(AB)_{ij}| = \left|\sum_{\nu} A_{i\nu}B_{\nu j}\right| \leq \sum_{\nu=1}^{n} |A_{i\nu}||B_{\nu j}| \leq n\|A\|\|B\|.$$

Thus $\|AB\| \leq n\|A\|\|B\|$. The second inequality follows by induction from the first inequality. □

Proof of Proposition (8.3). To prove that the matrix exponential converges absolutely, we estimate the series as follows: Let $a = n\|A\|$. Then

$$(8.23) \qquad |(e^A)_{ij}| \le |I_{ij}| + |A_{ij}| + \frac{1}{2!}|(A^2)_{ij}| + \frac{1}{3!}|(A^3)_{ij}| + \cdots$$
$$\le 1 + \|A\| + \frac{1}{2!}n\|A\|^2 + \frac{1}{3!}n^2\|A\|^3 + \cdots$$
$$= 1 + (a + \frac{1}{2!}a^2 + \frac{1}{3!}a^3 + \cdots)/n = 1 + (e^a - 1)/n. \square$$

Proof of Proposition (8.9).

(a) The terms of degree k in the expansions of (8.8) are

$$(x + y)^k/k! = \sum_{r+s=k}\binom{k}{r}x^r y^s/k! \quad \text{and} \quad \sum_{r+s=k}\frac{x^r}{r!}\frac{y^s}{s!}.$$

To show that these terms are equal, we have to show that

$$\binom{k}{r}/k! = \frac{1}{r!s!} \quad \text{or} \quad \binom{k}{r} = \frac{k!}{r!s!},$$

for all k and all r, s such that $r + s = k$. This is a standard formula for binomial coefficients.

(b) Denote by $S_n(x)$ the partial sum $1 + x/1! + x^2/2! + \cdots + x^n/n!$. Then

$$S_n(x)S_n(y) = (1 + x/1! + x^2/2! + \cdots + x^n/n!)(1 + y/1! + y^2/2! + \cdots + y^n/n!)$$
$$= \sum_{r,s=0}^{n}\frac{x^r}{r!}\frac{y^s}{s!},$$

while

$$S_n(x + y) = (1 + (x + y)/1! + (x + y)^2/2! + \cdots + (x + y)^n/n!)$$
$$= \sum_{k=0}^{n}\sum_{r+s=k}\binom{k}{r}x^r y^s/k! = \sum_{k=0}^{n}\sum_{r+s=k}\frac{x^r}{r!}\frac{y^s}{s!}.$$

Comparing terms, we find that the expansion of the partial sum $S_n(x + y)$ consists of the terms in $S_n(x)S_n(y)$ such that $r + s \le n$. The same is true when we substitute commuting matrices A, B for x, y. We must show that the sum of the remaining terms tends to zero as $k \longrightarrow \infty$.

(8.24) **Lemma.** The series $\sum_k \sum_{r+s=k}\left|\left(\frac{A^r}{r!}\frac{B^s}{s!}\right)_{ij}\right|$ converges for all i, j.

Proof. Let $a = n\|A\|$ and $b = n\|B\|$. We estimate the terms in the sum. According to (8.22), $|(A^rB^s)_{ij}| \le n(n^{r-1}\|A\|^r)(n^{s-1}\|B\|^s) \le a^r b^s$. Therefore

$$\sum_k \sum_{r+s=k}\left|\left(\frac{A^r}{r!}\frac{B^s}{s!}\right)_{ij}\right| \le \sum_k \sum_{r+s=k}\frac{a^r}{r!}\frac{b^s}{s!} = e^{a+b}.$$

The proposition follows from this lemma because, on the one hand, the i, j-entry of

$(S_k(A)S_k(B) - S_k(A+B))_{ij}$ is bounded by $\displaystyle\sum_{r+s>k}\left|\left(\frac{A^r}{r!}\frac{B^s}{s!}\right)_{ij}\right|$.

According to the lemma, this sum tends to zero as $k \longrightarrow \infty$. And on the other hand,

$$(S_k(A)S_k(B) - S_k(A+B)) \longrightarrow (e^A e^B - e^{A+B}). \square$$

Proof of Proposition (8.13). By definition,

$$\frac{d}{dt}(e^{tA}) = \lim_{h\to 0}\frac{e^{(t+h)A} - e^{tA}}{h}.$$

Since the matrices tA and hA commute, the Proposition (8.9) shows that

$$\frac{e^{(t+h)A} - e^{tA}}{h} = \left(\frac{e^{hA} - I}{h}\right)e^{tA}.$$

So our proposition follows from this lemma:

(8.25) Lemma. $\displaystyle\lim_{h\to 0}\frac{e^{hA} - I}{h} = A.$

Proof. The series expansion for the exponential shows that

$$\frac{e^{hA} - I}{h} - A = \frac{h}{2!}A^2 + \frac{h^2}{3!}A^3 + \cdots. \tag{8.26}$$

We estimate this series: Let $a = |h|n\|A\|$. Then

$$\left|\left(\frac{h}{2!}A^2 + \frac{h^2}{3!}A^3 + \cdots\right)_{ij}\right| \le \left|\frac{h}{2!}(A^2)_{ij}\right| + \left|\frac{h^2}{3!}(A^3)_{ij}\right| + \cdots$$

$$\le \frac{1}{2!}|h|n\|A\|^2 + \frac{1}{3!}|h|^2n^2\|A\|^3 + \cdots = \|A\|\left(\frac{1}{2!}a + \frac{1}{3!}a^2 + \cdots\right)$$

$$= \frac{\|A\|}{a}(e^a - 1 - a) = \|A\|\left(\frac{e^a - 1}{a} - 1\right).$$

Note that $a \longrightarrow 0$ as $h \longrightarrow 0$. Since the derivative of e^x is e^x,

$$\lim_{a\to 0}\frac{e^a - 1}{a} = \frac{d}{dx}e^x\bigg|_{x=0} = e^0 = 1.$$

So (8.26) tends to zero with h. $\square$

We will use the remarkable properties of the matrix exponential again, in Chapter 8.

I have not thought it necessary to undertake the labour of a formal proof of the theorem in the general case.

Arthur Cayley

EXERCISES

1. The Dimension Formula

1. Let T be left multiplication by the matrix $\begin{bmatrix} 1 & 2 & 0 & -1 & 5 \\ 2 & 0 & 2 & 0 & 1 \\ 1 & 1 & -1 & 3 & 2 \\ 0 & 3 & -3 & 2 & 6 \end{bmatrix}$. Compute $\ker T$ and $\operatorname{im} T$ explicitly by exhibiting bases for these spaces, and verify (1.7).
2. Determine the rank of the matrix $\begin{bmatrix} 11 & 12 & 13 & 14 \\ 21 & 22 & 23 & 24 \\ 31 & 32 & 33 & 34 \\ 41 & 42 & 43 & 44 \end{bmatrix}$.
3. Let $T: V \longrightarrow W$ be a linear transformation. Prove that $\ker T$ is a subspace of V and that $\operatorname{im} T$ is a subspace of W.
4. Let A be an $m \times n$ matrix. Prove that the space of solutions of the linear system $AX = 0$ has dimension at least $n - m$.
5. Let A be a $k \times m$ matrix and let B be an $n \times p$ matrix. Prove that the rule $M \rightsquigarrow AMB$ defines a linear transformation from the space $F^{m \times n}$ of $m \times n$ matrices to the space $F^{k \times p}$.
6. Let $(v_1, \dots, v_n)$ be a subset of a vector space V. Prove that the map $\varphi: F^n \longrightarrow V$ defined by $\varphi(X) = v_1x_1 + \cdots + v_nx_n$ is a linear transformation.
7. When the field is one of the fields $\mathbb{F}_p$, finite-dimensional vector spaces have finitely many elements. In this case, formula (1.6) and formula (6.15) from Chapter 2 both apply. Reconcile them.
8. Prove that every $m \times n$ matrix A of rank 1 has the form $A = XY^t$, where X, Y are m- and n-dimensional column vectors.
9. (a) The *left shift* operator S^- on $V = \mathbb{R}^\infty$ is defined by $(a_1, a_2, \dots,) \rightsquigarrow (a_2, a_3, \dots)$. Prove that $\ker S^- > 0$, but $\operatorname{im} S^- = V$.
   (b) The *right shift* operator S^+ on $V = \mathbb{R}^\infty$ is defined by $(a_1, a_2, \dots) \rightsquigarrow (0, a_1, a_2, \dots)$. Prove that $\ker S^+ = 0$, but $\operatorname{im} S^+ < V$.

2. The Matrix of a Linear Transformation

1. Determine the matrix of the differentiation operator $\frac{d}{dx}: P_n \longrightarrow P_{n-1}$ with respect to the natural bases (see (1.4)).
2. Find all linear transformations $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ which carry the line $y = x$ to the line $y = 3x$.
3. Prove Proposition (2.9b) using row and column operations.

4. Let T: $\mathbb{R}^3 \longrightarrow \mathbb{R}^2$ be the linear transformation defined by the rule $T(x_1, x_2, x_3)^t = (x_1 + x_2, 2x_3 - x_1)^t$. What is the matrix of T with respect to the standard bases?

5. Let A be an $n \times n$ matrix, and let $V = F^n$ denote the space of row vectors. What is the matrix of the linear operator "right multiplication by A" with respect to the standard basis of V?

6. Prove that different matrices define different linear transformations.

7. Describe left multiplication and right multiplication by the matrix (2.10), and prove that the rank of this matrix is r.

8. Prove that A and A^t have the same rank.

9. Let T_1, T_2 be linear transformations from V to W. Define $T_1 + T_2$ and cT by the rules $[T_1 + T_2](v) = T_1(v) + T_2(v)$ and $[cT](v) = cT(v)$.

(a) Prove that $T_1 + T_2$ and cT_1 are linear transformations, and describe their matrices in terms of the matrices for T_1, T_2.

(b) Let L be the set of all linear transformations from V to W. Prove that these laws make L into a vector space, and compute its dimension.

3. Linear Operators and Eigenvectors

1. Let V be the vector space of real 2×2 symmetric matrices $X = \begin{bmatrix} x & y \\ y & z \end{bmatrix}$, and let $A = \begin{bmatrix} 2 & 1 \\ & 1 \end{bmatrix}$. Determine the matrix of the linear operator on V defined by $X \rightsquigarrow AXA^t$, with respect to a suitable basis.

2. Let $A = (a_{ij})$, $B = (b_{ij})$ be 2×2 matrices, and consider the operator T: $M \rightsquigarrow AMB$ on the space $F^{2 \times 2}$ of 2×2 matrices. Find the matrix of T with respect to the basis $(e_{11}, e_{12}, e_{21}, e_{22})$ of $F^{2 \times 2}$.

3. Let T: $V \longrightarrow V$ be a linear operator on a vector space of dimension 2. Assume that T is not multiplication by a scalar. Prove that there is a vector $v \in V$ such that $(v, T(v))$ is a basis of V, and describe the matrix of T with respect to that basis.

4. Let T be a linear operator on a vector space V, and let $c \in F$. Let W be the set of eigenvectors of T with eigenvalue c, together with 0. Prove that W is a T-invariant subspace.

5. Find all invariant subspaces of the real linear operator whose matrix is as follows.

(a) $\begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & & \\ & 2 & \\ & & 3 \end{bmatrix}$

6. An operator on a vector space V is called *nilpotent* if $T^k = 0$ for some k. Let T be a nilpotent operator, and let $W^i = \text{im } T^i$.

(a) Prove that if $W^i \neq 0$, then $\dim W^{i+1} < \dim W^i$.

(b) Prove that if V is a space of dimension n and if T is nilpotent, then $T^n = 0$.

7. Let T be a linear operator on $\mathbb{R}^2$. Prove that if T carries a line ℓ to ℓ, then it also carries every line parallel to l to another line parallel to l.

8. Prove that the composition $T_1 \circ T_2$ of linear operators on a vector space is a linear operator, and compute its matrix in terms of the matrices A_1, A_2 of T_1, T_2.

9. Let P be the real vector space of polynomials $p(x) = a_0 + a + \cdots + a_n x^n$ of degree $\leq n$, and let D denote the derivative $\frac{d}{dx}$, considered as a linear operator on P.

(a) Find the matrix of D with respect to a convenient basis, and prove that D is a nilpotent operator.
(b) Determine all the D-invariant subspaces.

10. Prove that the matrices $\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$ and $\begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$ $(b \neq 0)$ are similar if and only if $a \neq d$.

11. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a real 2×2 matrix. Prove that A can be reduced to a matrix $\begin{bmatrix} 0 & * \\ 1 & * \end{bmatrix}$ by row and column operations of the form $A \longrightarrow EAE^{-1}$, unless $b = c = 0$ and $a = d$. Make a careful case analysis to take care of the possibility that b or c is zero.

12. Let T be a linear operator on $\mathbb{R}^2$ with two linearly independent eigenvectors v_1, v_2. Assume that the eigenvalues c_1, c_2 of these operators are positive and that $c_1 > c_2$. Let ℓ_i be the line spanned by v_i.
(a) The operator T carries every line ℓ through the origin to another line. Using the parallelogram law for vector addition, show that every line $\ell \neq \ell_2$ is shifted away from ℓ_2 toward ℓ_1.
(b) Use (a) to prove that the only eigenvectors are multiples of v_1 or v_2.
(c) Describe the effect on lines when there is a single line carried to itself, with positive eigenvalue.

13. Consider an arbitrary 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. The condition that a column vector X be an eigenvector for left multiplication by A is that $Y = AX$ be parallel to X, which means that the slopes $s = x_2/x_1$ and $s' = y_2/y_1$ are equal.
(a) Find the equation in s which expresses this equality.
(b) For which A is $s = 0$ a solution? $s = \infty$?
(c) Prove that if the entries of A are positive real numbers, then there is an eigenvector in the first quadrant and also one in the second quadrant.

4. The Characteristic Polynomial

1. Compute the characteristic polynomials, eigenvalues, and eigenvectors of the following complex matrices.
(a) $\begin{bmatrix} -2 & 2 \\ -2 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$

2. (a) Prove that the eigenvalues of a real symmetric 2×2 matrix are real numbers.
(b) Prove that a real 2×2 matrix whose off-diagonal entries are positive has real eigenvalues.

3. Find the complex eigenvalues and eigenvectors of the notation matrix

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

4. Prove that a real 3×3 matrix has at least one real eigenvalue.

5. Determine the characteristic polynomial of the matrix

$$\begin{bmatrix} 0 & 1 & & & \\ 1 & 0 & 1 & & \\ & 1 & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & 1 & 0 \end{bmatrix}.$$

6. Prove Proposition (4.18).
7. (a) Let T be a linear operator having two linearly independent eigenvectors with the same eigenvalue λ. Is it true that λ is a multiple root of the characteristic polynomial of T?
   (b) Suppose that λ is a multiple root of the characteristic polynomial. Does T have two linearly independent eigenvectors with eigenvalue λ?
8. Let V be a vector space with basis $(v_1,\ldots,v_n)$ over a field F, and let $a_1,\ldots,a_{n-1}$ be elements of F. Define a linear operator on V by the rules $T(v_i) = v_{i+1}$ if $i < n$ and $T(v_n) = a_1v_1 + a_2v_2 + \cdots + a_{n-1}v_{n-1}$.
   (a) Determine the matrix of T with respect to the given basis.
   (b) Determine the characteristic polynomial of T.
9. Do A and A^t have the same eigenvalues? the same eigenvectors?
10. (a) Use the characteristic polynomial to prove that a 2×2 real matrix P all of whose entries are positive has two distinct real eigenvalues.
    (b) Prove that the larger eigenvalue has an eigenvector in the first quadrant, and the smaller eigenvalue has an eigenvector in the second quadrant.
11. (a) Let A be a 3×3 matrix, with characteristic polynomial

$$p(t) = t^3 - (\operatorname{tr} A)t^2 + s_1 t - (\det A).$$

Prove that s_1 is the sum of the symmetric 2×2 subdeterminants:

$$s_1 = \det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \det\begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix} + \det\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}.$$

   *(b) Generalize to $n \times n$ matrices.
12. Let T be a linear operator on a space of dimension n, with eigenvalues $\lambda_1,\ldots,\lambda_n$.
    (a) Prove that $\operatorname{tr} T = \lambda_1 + \cdots + \lambda_n$ and that $\det T = \lambda_1 \cdots \lambda_n$.
    (b) Determine the other coefficients of the characteristic polynomial in terms of the eigenvalues.
*13. Consider the linear operator of left multiplication of an $n \times n$ matrix A on the space $F^{n\times n}$ of all $n \times n$ matrices. Compute the trace and the determinant of this operator.
*14. Let P be a real matrix such that $P^t = P^2$. What are the possible eigenvalues of P?
15. Let A be a matrix such that $A^n = I$. Prove that the eigenvalues of A are powers of nth root of unity $\zeta_n = e^{2\pi i/n}$.

5. Orthogonal Matrices and Rotations

1. What is the matrix of the three-dimensional rotation through the angle θ about the axis e_2?
2. Prove that every orthonormal set of n vectors in $\mathbb{R}^n$ is a basis.
3. Prove algebraically that a real 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ represents a rotation if and only if it is in SO_2.
4. (a) Prove that O_n and SO_n are subgroups of $GL_n(\mathbb{R})$, and determine the index of SO_n in O_n.
   (b) Is O_2 isomorphic to the product group $SO_2 \times \{\pm I\}$? Is O_3 isomorphic to $SO_3 \times \{\pm I\}$?

5. What are the eigenvalues of the matrix A which represents the rotation of $\mathbb{R}^3$ by θ about an axis v?
6. Let A be a matrix in O_3 whose determinant is -1. Prove that -1 is an eigenvalue of A.
7. Let A be an orthogonal 2×2 matrix whose determinant is -1. Prove that A represents a reflection about a line through the origin.
8. Let A be an element of SO_3, with angle of rotation θ. Show that $\cos\theta = \frac{1}{2}(\operatorname{tr} A - 1)$.
9. Every real polynomial of degree 3 has a real root. Use this fact to give a less tricky proof of Lemma (5.23).

*10. Find a geometric way to determine the axis of rotation for the composition of two three-dimensional rotations.

11. Let v be a vector of unit length, and let P be the plane in $\mathbb{R}^3$ orthogonal to v. Describe a bijective correspondence between points on the unit circle in P and matrices $P \in SO_3$ whose first column is v.
12. Describe geometrically the action of an orthogonal matrix with determinant -1.
13. Prove that a rigid motion, as defined by (5.15), is bijective.

*14. Let A be an element of SO_3. Show that if it is defined, the vector

$$((a_{23} + a_{32})^{-1}, (a_{13} + a_{31})^{-1}, (a_{12} + a_{21})^{-1})^t$$

is an eigenvector with eigenvalue 1.

6. Diagonalization

1. (a) Find the eigenvectors and eigenvalues of the matrix

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

   (b) Find a matrix P such that PAP^{-1} is diagonal.

   (c) Compute $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}^{30}$.

2. Diagonalize the rotation matrix $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, using complex numbers.
3. Prove that if A, B are $n \times n$ matrices and if A is nonsingular, then AB is similar to BA.
4. Let A be a complex matrix having zero as its only eigenvalue. Prove or disprove: A is nilpotent.
5. In each case, if the matrix is diagonalizable, find a matrix P such that PAP^{-1} is diagonal.

   (a) $\begin{bmatrix} -2 & 2 \\ -2 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

6. Can the diagonalization (6.1) be done with a matrix $P \in SL_n$?
7. Prove that a linear operator T is nilpotent if and only if there is a basis of V such that the matrix of T is upper triangular, with diagonal entries zero.
8. Let T be a linear operator on a space of dimension 2. Assume that the characteristic polynomial of T is $(t - a)^2$. Prove that there is a basis of V such that the matrix of T has one of the two forms $\begin{bmatrix} a & 1 \\ 0 & a \end{bmatrix}, \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$.

9. Let A be a nilpotent matrix. Prove that $\det(I + A) = 1$.

10. Prove that if A is a nilpotent $n \times n$ matrix, then $A^n = 0$.

11. Find all real 2×2 matrices such that $A^2 = I$, and describe geometrically the way they operate by left multiplication on $\mathbb{R}^2$.

12. Let M be a matrix made up of two diagonal blocks: $M = \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}$. Prove that M is diagonalizable if and only if A and D are.

13. **(a)** Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a 2×2 matrix with eigenvalue λ. Show that $(b, \lambda - a)^t$ is an eigenvector for A.

(b) Find a matrix P such that PAP^{-1} is diagonal, if A has two distinct eigenvalues $\lambda_1 \neq \lambda_2$.

14. Let A be a complex $n \times n$ matrix. Prove that there is a matrix B arbitrarily close to A (meaning that $|b_{ij} - a_{ij}|$ can be made arbitrarily small for all i, j) such that B has n distinct eigenvalues.

***15.** Let A be a complex $n \times n$ matrix with n distinct eigenvalues $\lambda_1, \ldots, \lambda_n$. Assume that λ_1 is the largest eigenvalue, that is, that $|\lambda_1| > |\lambda_i|$ for all $i > 1$. Prove that for most vectors X the sequence $X_k = \lambda_1^{-k}A^kX$ converges to an eigenvector Y with eigenvalue λ_1, and describe precisely what the conditions on X are for this to be the case.

16. **(a)** Use the method of the previous problem to compute the largest eigenvalue of the matrix $\begin{bmatrix} 3 & 1 \\ 3 & 4 \end{bmatrix}$ to three-place accuracy.

(b) Compute the largest eigenvalue of the matrix $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$ to three-place accuracy.

***17.** Let A be $m \times m$ and B be $n \times n$ complex matrices, and consider the linear operator T on the space $F^{m \times n}$ of all complex matrices defined by $T(M) = AMB$.

(a) Show how to construct an eigenvector for T out of a pair of column vectors X, Y, where X is an eigenvector for A and Y is an eigenvector for B^t.

(b) Determine the eigenvalues of T in terms of those of A and B.

***18.** Let A be an $n \times n$ complex matrix.

(a) Consider the linear operator T defined on the space $F^{n \times n}$ of all complex $n \times n$ matrices by the rule $T(B) = AB - BA$. Prove that the rank of this operator is at most $n^2 - n$.

(b) Determine the eigenvalues of T in terms of the eigenvalues $\lambda_1, \ldots, \lambda_n$ of A.

7. Systems of Differential Equations

1. Let v be an eigenvector for the matrix A, with eigenvalue c. Prove that $e^{ct}v$ solves the differential equation $\frac{dX}{dt} = AX$.

2. Solve the equation $\frac{dX}{dt} = AX$ for the following matrices A:

(a) $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ **(b)** $\begin{bmatrix} -2 & 2 \\ -2 & 3 \end{bmatrix}$ **(c)** $\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$ **(d)** $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$ **(e)** $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

3. Explain why diagonalization gives the general solution.

4. (a) Prove Proposition (7.16).
   (b) Why is it enough to write down the real and imaginary parts to get the general solution?
5. Prove Lemma (7.25).
6. Solve the inhomogeneous differential equation $\frac{dX}{dt} = AX + B$ in terms of the solutions to the homogeneous equation $\frac{dX}{dt} = AX$.
7. A differential equation of the form $d^nx/dt^n + a_{n-1}d^{n-1}x/dt^{n-1} + \cdots + a_1dx/dt + a_0x = 0$ can be rewritten as a system of first-order equations by the following trick: We introduce unknown functions $x_0, x_1, \ldots, x_{n-1}$ with $x = x_0$, and we set $dx_i/dt = x_{i+1}$ for $i = 0, \ldots, n-2$. The original equation can be rewritten as the system $dx_i/dt = x_{i+1}$, $i = 0, \ldots, n-2$, and $dx_{n-1}/dt = -(a_{n-1}x_{n-1} + \cdots + a_1x_1 + a_0x)$. Determine the matrix which represents this system of equations.
8. (a) Rewrite the second-order linear equation in one variable

$$\frac{d^2x}{dt^2} + b\frac{dx}{dt} + cx = 0$$

   as a system of two first-order equations in two unknowns $x_0 = x$, $x_1 = dx/dt$.
   (b) Solve the system when $b = -4$ and $c = 3$.
9. Let A be an $n \times n$ matrix, and let $B(t)$ be a column vector of continuous functions on the interval $[\alpha, \beta]$. Define $F(t) = \int_\alpha^t e^{-tA}B(t)\,dt$.
   (a) Prove that $X = F(t)$ is a solution of the differential equation $X' = AX + B(t)$ on the interval (α, β).
   (b) Determine all solutions of this equation on the interval.

8. The Matrix Exponential

1. Compute e^A for the following matrices A:
   (a) $\begin{bmatrix} & 1 \\ -1 & \end{bmatrix}$ (b) $\begin{bmatrix} a & b \\ & \end{bmatrix}$
2. Let $A = \begin{bmatrix} 1 & 1 \\ & 2 \end{bmatrix}$.
   (a) Compute e^A directly from the expansion.
   (b) Compute e^A by diagonalizing the matrix.
3. Compute e^A for the following matrices A:
   (a) $\begin{bmatrix} 0 & -b \\ b & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}$
4. Compute e^A for the following matrices A:
   (a) $\begin{bmatrix} 2\pi i & 2\pi i \\ & 2\pi i \end{bmatrix}$ (b) $\begin{bmatrix} 6\pi i & 4\pi i \\ 2\pi i & 8\pi i \end{bmatrix}$
5. Let A be an $n \times n$ matrix. Prove that the map $t \rightsquigarrow e^{tA}$ is a homomorphism from the additive group $\mathbb{R}^+$ to $GL_n(\mathbb{C})$.

6. Find two matrices A, B such that $e^{A+B} \neq e^A e^B$.
7. Prove the formula $e^{\text{trace } A} = \det(e^A)$.
8. Solve the differential equation $\frac{dX}{dt} = AX$, when $A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$.
9. Let $f(t)$ be a polynomial, and let T be a linear operator. Prove that $f(T)$ is a linear operator.
10. Let A be a symmetric matrix, and let $f(t)$ be a polynomial. Prove that $f(A)$ is symmetric.
11. Prove the product rule for differentiation of matrix-valued functions.
12. Let $A(t)$, $B(t)$ be differentiable matrix-valued functions of t. Compute the following.
    (a) $d/dt\,(A(t)^3)$
    (b) $d/dt(A(t)^{-1})$, assuming that $A(t)$ is invertible for all t
    (c) $d/dt(A(t)^{-1}B(t))$
13. Let X be an eigenvector of an $n \times n$ matrix A, with eigenvalue λ.
    (a) Prove that if A is invertible then X is also an eigenvector for A^{-1}, and that its eigenvalue is λ^{-1}.
    (b) Let $p(t)$ be a polynomial. Then X is an eigenvector for $p(A)$, with eigenvalue $p(\lambda)$.
    (c) Prove that X is an eigenvector for e^A, with eigenvalue e^λ.
14. For an $n \times n$ matrix A, define $\sin A$ and $\cos A$ by using the Taylor's series expansions for $\sin x$ and $\cos x$.
    (a) Prove that these series converge for all A.
    (b) Prove that $\sin tA$ is a differentiable function of t and that $d(\sin tA)/dt = A \cos tA$.
15. Discuss the range of validity of the following identities.
    (a) $\cos^2 A + \sin^2 A = I$
    (b) $e^{iA} = \cos A + i \sin A$
    (c) $\sin(A + B) = \sin A \cos B + \cos A \sin B$
    (d) $\cos(A + B) = \cos A \cos B - \sin A \sin B$
    (e) $e^{2\pi i A} = I$
    (f) $d(e^{A(t)})/dt = e^{A(t)} A'(t)$, where $A(t)$ is a differentiable matrix-valued function of t.
16. (a) Derive the product rule for differentiation of complex-valued functions in two ways: directly, and by writing $x(t) = u(t) + iv(t)$ and applying the product rule for real-valued functions.
    (b) Let $f(t)$ be a complex-valued function of a real variable t, and let $\varphi(u)$ be a real-valued function of u. State and prove the chain rule for $f(\varphi(u))$.
17. (a) Let B_k be a sequence of $m \times n$ matrices which converges to a matrix B, and let P be an $m \times m$ matrix. Prove that PB_k converges to PB.
    (b) Prove that if $m = n$ and P is invertible, then PB_kP^{-1} converges to PBP^{-1}.
18. Let $f(x) = \Sigma c_k x^k$ be a power series such that $\Sigma c_k A^k$ converges when A is a sufficiently small $n \times n$ matrix. Prove that A and $f(A)$ commute.
19. Determine $\frac{d}{dt} \det A(t)$, when $A(t)$ is a differentiable matrix function of t.

Miscellaneous Problems

1. What are the possible eigenvalues of a linear operator T such that (a) $T^r = I$, (b) $T^r = 0$, (c) $T^2 - 5T + 6 = 0$?

2. A linear operator T is called nilpotent if some power of T is zero.
   (a) Prove that T is nilpotent if and only if its characteristic polynomial is t^n, $n = \dim V$.
   (b) Prove that if T is a nilpotent operator on a vector space of dimension n, then $T^n = 0$.
   (c) A linear operator T is called *unipotent* if $T - I$ is nilpotent. Determine the characteristic polynomial of a unipotent operator. What are its possible eigenvalues?
3. Let A be an $n \times n$ complex matrix. Prove that if trace $A^i = 0$ for all i, then A is nilpotent.
*4. Let A, B be complex $n \times n$ matrices, and let $C = AB - BA$. Prove that if C commutes with A then C is nilpotent.
5. Let $\lambda_1, \dots, \lambda_n$ be the roots of the characteristic polynomial $p(t)$ of a complex matrix A. Prove the formulas trace $A = \lambda_1 + \cdots + \lambda_n$ and $\det A = \lambda_1 \cdots \lambda_n$.
6. Let T be a linear operator on a real vector space V such that $T^2 = I$. Define subspaces as follows:

$$W^+ = \{v \in V \mid T(v) = v\}, \quad W^- = \{v \in V \mid T(v) = -v\}.$$

   Prove that V is isomorphic to the direct sum $W^+ \oplus W^-$.
7. The *Frobenius norm* $|A|$ of an $n \times n$ matrix A is defined to be the length of A when it is considered as an n^2-dimensional vector: $|A|^2 = \Sigma\, |a_{ij}|^2$. Prove the following inequalities: $|A + B| \le |A| + |B|$ and $|AB| \le |A|\,|B|$.
8. Let $T: V \longrightarrow V$ be a linear operator on a finite-dimensional vector space V. Prove that there is an integer n so that $(\ker T^n) \cap (\operatorname{im} T^n) = 0$.
9. Which infinite matrices represent linear operators on the space Z [Chapter 3 (5.2d)]?
*10. The $k \times k$ *minors* of an $m \times n$ matrix A are the square submatrices obtained by crossing out $m - k$ rows and $n - k$ columns. Let A be a matrix of rank r. Prove that some $r \times r$ minor is invertible and that no $(r + 1) \times (r + 1)$ minor is invertible.
11. Let $\varphi: F^n \longrightarrow F^m$ be left multiplication by an $m \times n$ matrix A. Prove that the following are equivalent.
   (a) A has a right inverse, a matrix B such that $AB = I$.
   (b) φ is surjective.
   (c) There is an $m \times m$ minor of A whose determinant is not zero.
12. Let $\varphi: F^n \longrightarrow F^m$ be left multiplication by an $m \times n$ matrix A. Prove that the following are equivalent.
   (a) A has a left inverse, a matrix B such that $BA = I$.
   (b) φ is injective.
   (c) There is an $n \times n$ *minor* of A whose determinant is not zero.
*13. Let A be an $n \times n$ matrix such that $A^r = I$. Prove that if A has only one eigenvalue ζ, then $A = \zeta I$.
14. (a) Without using the characteristic polynomial, prove that a linear operator on a vector space of dimension n can have at most n different eigenvalues.
   (b) Use (a) to prove that a polynomial of degree n with coefficients in a field F has at most n roots in F.
15. Let A be an $n \times n$ matrix, and let $p(t) = t^n + c_{n-1}t^{n-1} + \cdots + c_1 t + c_0$ be its characteristic polynomial. The *Cayley–Hamilton Theorem* asserts that

$$p(A) = A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 I = 0.$$

   (a) Prove the Cayley–Hamilton Theorem for 2×2 matrices.
   (b) Prove it for diagonal matrices.

(c) Prove it for diagonalizable matrices.

*(d) Show that every complex $n \times n$ matrix is arbitrarily close to a diagonalizable matrix, and use this fact to extend the proof for diagonalizable matrices to all complex matrices by continuity.

16. (a) Use the Cayley–Hamilton Theorem to give an expression for A^{-1} in terms of A, $(\det A)^{-1}$, and the coefficients of the characteristic polynomial.

(b) Verify this expression in the 2×2 case by direct computation.

***17.** Let A be a 2×2 matrix. The Cayley–Hamilton Theorem allows all powers of A to be written as linear combinations of I and A. Therefore it is plausible that e^A is also such a linear combination.

(a) Prove that if a, b are the eigenvalues of A and if $a \neq b$, then

$$e^A = \frac{ae^b - be^a}{a - b} I + \frac{e^a - e^b}{a - b} A.$$

(b) Find the correct formula for the case that A has two equal eigenvalues.

18. The Fibonacci numbers $0, 1, 1, 2, 3, 5, 8, \ldots$ are defined by the recursive relations $f_n = f_{n-1} + f_{n-2}$, with the initial conditions $f_0 = 0, f_1 = 1$. This recursive relation can be written in matrix form as

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} f_{n-2} \\ f_{n-1} \end{bmatrix} = \begin{bmatrix} f_{n-1} \\ f_n \end{bmatrix}.$$

(a) Prove the formula

$$f_n = \frac{1}{\alpha}\left[\left(\frac{1 + \alpha}{2}\right)^n - \left(\frac{1 - \alpha}{2}\right)^n\right],$$

where $\alpha = \sqrt{5}$.

(b) Suppose that the sequence a_n is defined by the relation $a_n = \frac{1}{2}(a_{n-1} + a_{n-2})$. Compute $\lim a_n$ in terms of a_0, a_1.

***19.** Let A be an $n \times n$ real positive matrix, and let $X \in \mathbb{R}^n$ be a column vector. Let us use the shorthand notation $X > 0$ or $X \geq 0$ to mean that all entries of the vector X are positive or nonnegative, respectively. By "positive quadrant" we mean the set of vectors $X \geq 0$. (But note that $X \geq 0$ and $X \neq 0$ do not imply $X > 0$ in our sense.)

(a) Prove that if $X \geq 0$ and $X \neq 0$ then $AX > 0$.

(b) Let C denote the set of pairs (X, t), $t \in \mathbb{R}$, such that $X \geq 0$, $|X| = 1$, and $(A - tI)X \geq 0$. Prove that C is a compact set in $\mathbb{R}^{n+1}$.

(c) The function t takes on a maximum value on C, say at the point (X_0, t_0). Then $(A - t_0 I)X_0 \geq 0$. Prove that $(A - t_0 I)X_0 = 0$.

(d) Prove that X_0 is an eigenvector with eigenvalue t_0 by showing that otherwise the vector $AX_0 = X_1$ would contradict the maximality of t_0.

(e) Prove that t_0 is the eigenvalue of A with largest absolute value.

***20.** Let $A = A(t)$ be a matrix of functions. What goes wrong when you try to prove that, in analogy with $n = 1$, the matrix

$$\exp\left(\int_{t_0}^{t} A(u)du\right)$$

is a solution of the system $dX/dt = AX$? Can you find conditions on the matrix function $A(t)$ which will make this a solution?

Chapter 5

Symmetry

L'algèbre n'est qu'une géométrie écrite;
la géométrie n'est qu'une algèbre figurée.

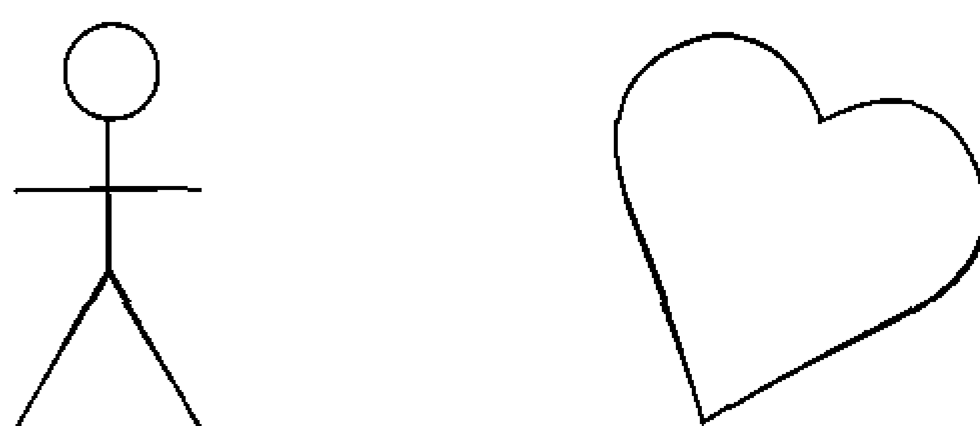
Sophie Germain

The study of symmetry provides one of the most appealing applications of group theory. Groups were first invented to analyze symmetries of certain algebraic structures called field extensions, and because symmetry is a common phenomenon in all sciences, it is still one of the two main ways in which group theory is applied. The other way is through group representations, which will be discussed in Chapter 9. In the first four sections of this chapter, we will study the symmetry of plane figures in terms of groups of rigid motions of the plane. Plane figures provide a rich source of examples and a background for the general concept of group operation, which is introduced in Section 5.

When studying symmetry, we will allow ourselves to use geometric reasoning without bothering to carry the arguments back to the axioms of geometry. That can be left for another occasion.

1. SYMMETRY OF PLANE FIGURES

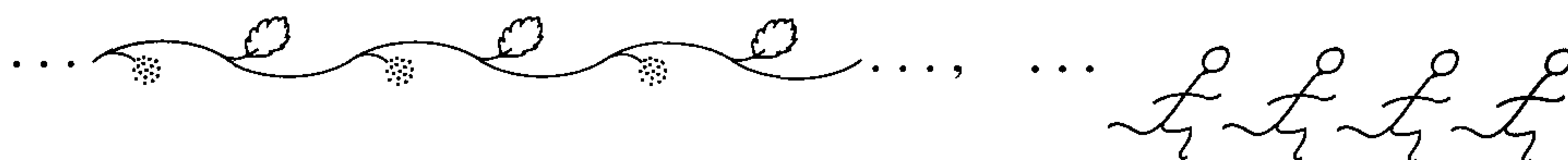
The possible symmetry of plane figures is usually classified into the main types shown in Figures (1.1–1.3).



(1.1) **Figure.** Bilateral symmetry.

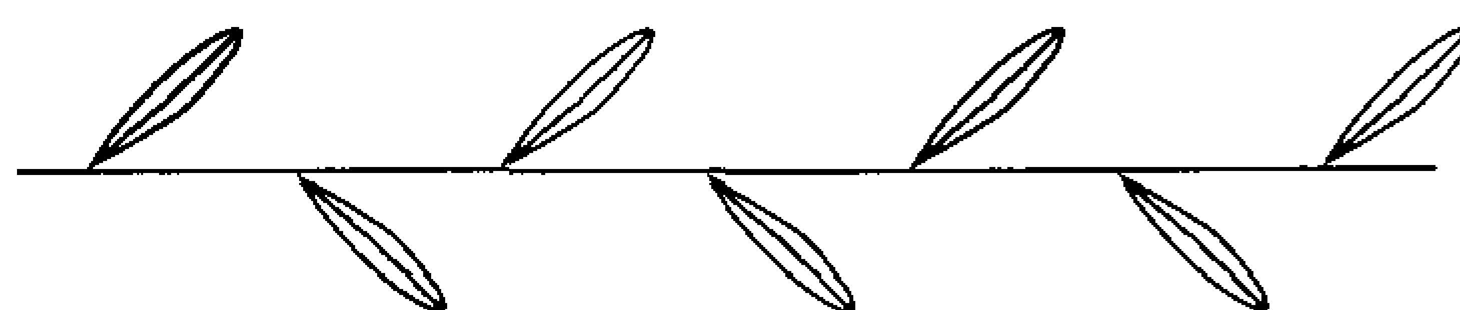


(1.2) **Figure.** Rotational symmetry.



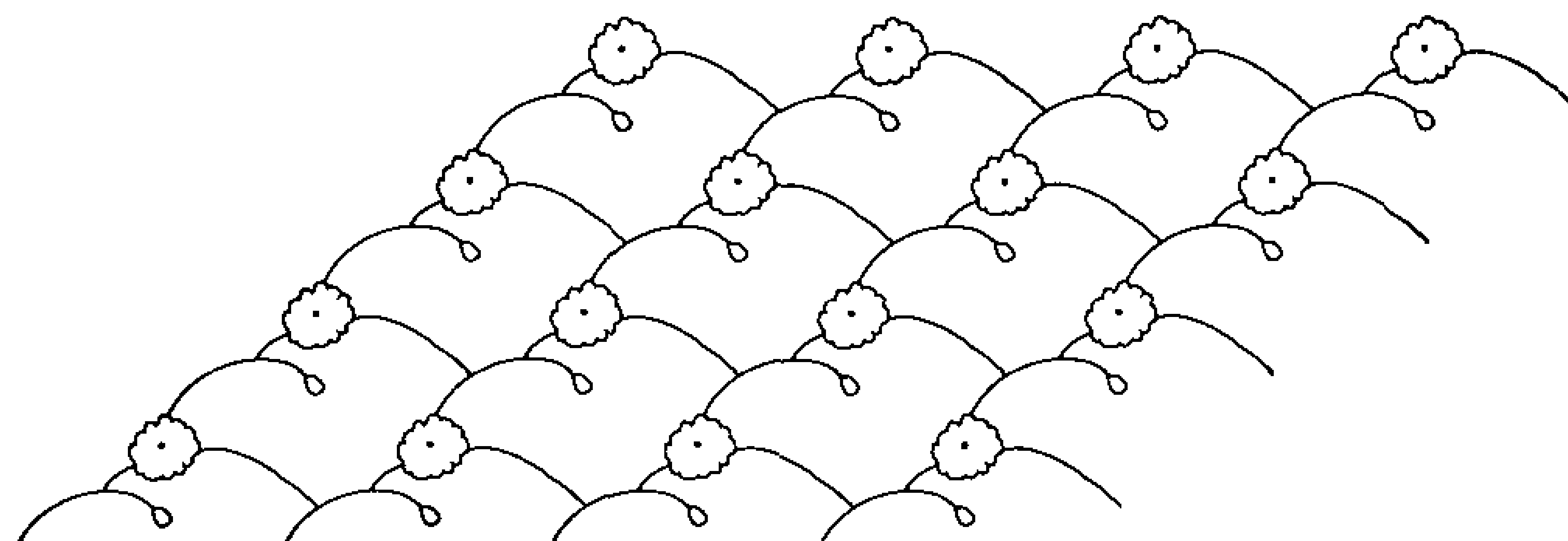
(1.3) **Figure.** Translational symmetry.

A fourth type of symmetry also exists, though it may be slightly less familiar:



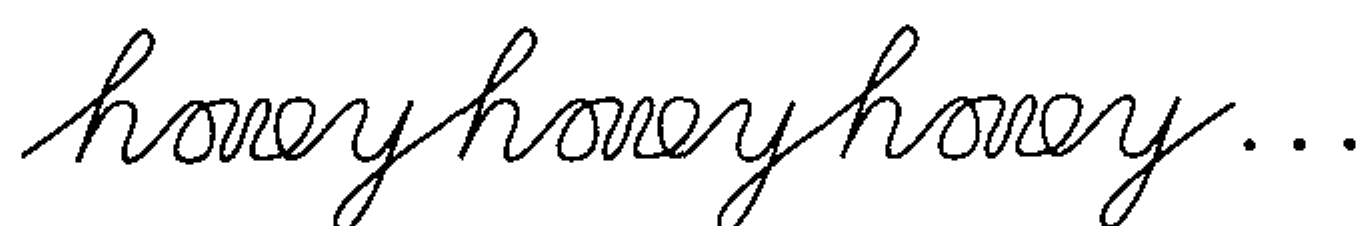
(1.4) **Figure.** Glide symmetry.

Figures such as wallpaper patterns may have two independent translational symmetries, as shown in Figure (1.5):



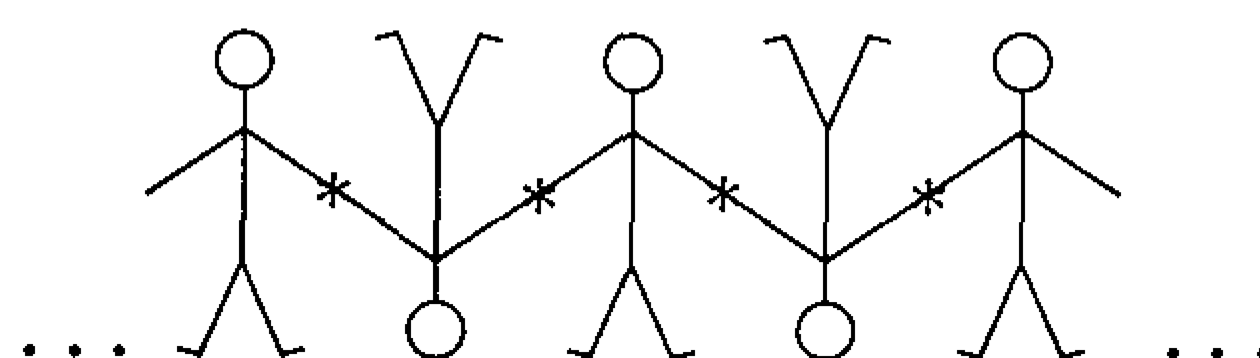
(1.5) **Figure.**

Other combinations of symmetries may also occur. For instance, the star has bilateral as well as rotational symmetry. Figure (1.6) is an example in which translational and rotational symmetry are combined:



(1.6) **Figure.**

Another example is shown in Figure (1.7).



(1.7) **Figure.**

As in Section 5 of Chapter 4, we call a map $m: P \longrightarrow P$ from the plane P to itself a *rigid motion,* or an *isometry,* if it is distance-preserving, that is, if for any two points $p, q \in P$ the distance from p to q is equal to the distance from $m(p)$ to $m(q)$. We will show in the next section that the rigid motions are translations, rotations, reflections, and glide reflections. They form a group M whose law of composition is composition of functions.

If a rigid motion m carries a subset F of the plane to itself, we call it a *symmetry* of F. The set of all symmetries of F always forms a subgroup G of M, called the *group of symmetries* of the figure. The fact that G is a subgroup is clear: If m and m' carry F to F, then so does the composed map mm', and so on.

The group of symmetries of the bilaterally symmetric Figure (1.1) consists of two elements: the identity transformation 1 and the reflection r about a line called the axis of symmetry. We have the relation $rr = 1$, which shows that G is a cyclic group of order 2, as it must be, because there is no other group of order 2.

The group of symmetries of Figure (1.3) is an infinite cyclic group generated by the motion which carries it one unit to the left. We call such a motion a *translation* t:

$$G = \{\dots, t^{-2}, t^{-1}, 1, t, t^2, \dots\}.$$

The symmetry groups of Figures (1.4, 1.6, 1.7) contain elements besides translations and are therefore larger. Do the exercise of describing their elements.

2. THE GROUP OF MOTIONS OF THE PLANE

This section describes the group M of all rigid motions of the plane. The coarsest classification of motions is into the *orientation-preserving* motions, those which do not flip the plane over, and the *orientation-reversing* motions which do flip it over (see Chapter 4, Section 5). We can use this partition of M to define a map

$$M \longrightarrow \{\pm 1\},$$

by sending the orientation-preserving motions to 1 and the orientation-reversing motions to -1. You will convince yourself without difficulty that this map is a homomorphism: The product of two orientation-reversing motions is orientation-preserving, and so on.

A finer classification of the motions is as follows:

(2.1)

(a) *The orientation-preserving motions:*
  (i) *Translation:* parallel motion of the plane by a vector a: $p \rightsquigarrow p+a$.
  (ii) *Rotation:* rotates the plane by an angle $\theta \neq 0$ about some point.

(b) *The orientation-reversing motions:*
  (i) *Reflection* about a line ℓ.
  (ii) *Glide reflection:* obtained by reflecting about a line ℓ, and then translating by a nonzero vector a parallel to ℓ.

(2.2) **Theorem.** The above list is complete. Every rigid motion is a translation, a rotation, a reflection, a glide reflection, or the identity.

This theorem is remarkable. One consequence is that the composition of rotations about two different points is a rotation about a third point, unless it is a translation. This fact follows from the theorem, because the composition preserves orientation, but it is not obvious.

Some of the other compositions are easier to visualize. The composition of rotations through angles θ and η about the same point is again a rotation, through the angle $\theta + \eta$, about that point. The composition of translations by the vectors a and b is the translation by their sum $a + b$.

Note that a translation does not leave any point fixed (unless the vector a is zero, in which case it is the identity map). Glides do not have fixed points either. On the other hand, a rotation fixes exactly one point, the center of rotation, and a reflection fixes the points on the line of reflection. Hence the composition of reflections about two nonparallel lines ℓ_1, ℓ_2 is a rotation about the intersection point $p = \ell_1 \cap \ell_2$. This follows from the theorem, because the composition does fix p, and it is orientation-preserving. The composition of two reflections about parallel lines is a translation by a vector orthogonal to the lines.

In order to prove Theorem (2.2), and also to be able to compute conveniently in the group M, we are going to choose some special motions as generators for the group. We will obtain defining relations similar to the relations (1.18) in Chapter 2 which define the symmetric group S_3, but since M is infinite, there will be more of them.

Let us identify the plane with the space $\mathbb{R}^2$ of column vectors, by choosing a coordinate system. Having done this, we choose as generators the translations, the rotations about the origin, and the reflection about the x_1-axis:

(2.3)

(a) *Translation* t_a by a vector a: $t_a(x) = x + a = \begin{bmatrix} x_1 + a_1 \\ x_2 + a_2 \end{bmatrix}$.

(b) *Rotation* ρ_θ by an angle θ about the origin:

$$\rho_\theta(x) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

(c) *Reflection* r about the x_1-axis: $r(x) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ -x_2 \end{bmatrix}$.

Since they fix the origin, the rotations ρ_θ and the reflection r are orthogonal operators on $\mathbb{R}^2$. A translation is not a linear operator—it does not send zero to itself, except of course for translation by the zero vector.

The motions (2.3) are not all of the elements of M. For example, rotation about a point other than the origin is not listed, nor are reflections about other lines.

However, they do generate the group: Every element of M is a product of such elements. It is easily seen that any rigid motion m can be obtained by composing them. Either

$$m = t_a\rho_\theta \quad \text{or else} \quad m = t_a\rho_\theta r, \tag{2.4}$$

for some vector a and angle θ, possibly zero. To see this, we recall that every rigid motion is the composition of an orthogonal operator followed by a translation [Chapter 4 (5.20)]. So we can write m in the form $m = t_a m'$, where m' is an orthogonal operator. Next, if $\det m' = 1$, then it is one of the rotations ρ_θ. This follows from Theorem (5.5) of Chapter 4. So in this case, $m = t_a\rho_\theta$. Finally, if $\det m' = -1$, then $\det m'r = 1$, so $m'r$ is a rotation ρ_θ. Since $r^2 = 1$, $m' = \rho_\theta r$ in this case, and $m = t_a\rho_\theta r$.

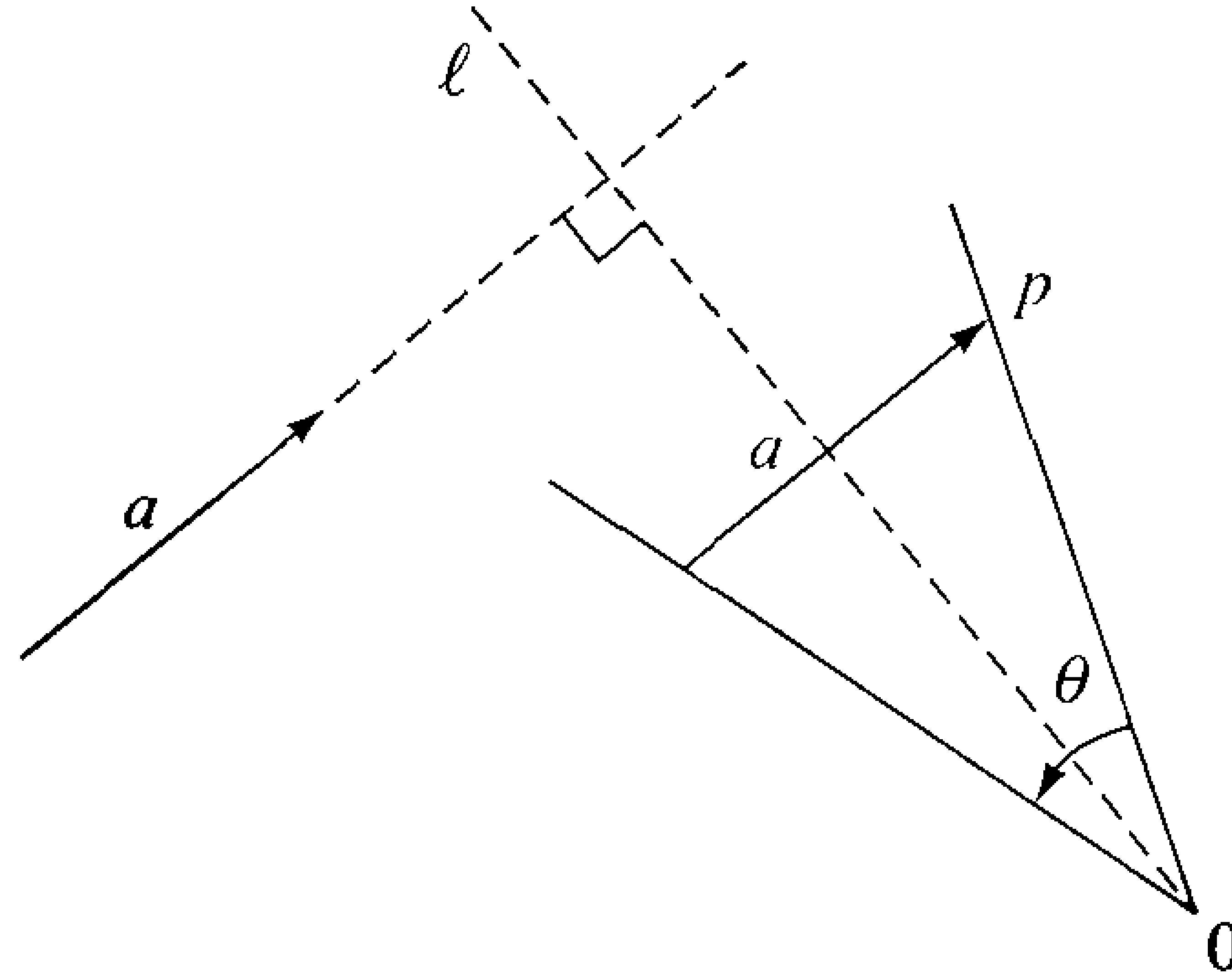
The expression of a motion m as a product (2.4) is unique. For suppose that m is expressed in two ways: $m = t_a\rho_\theta r^i = t_b\rho_\eta r^j$, where i, j are 0 or 1. Since m is orientation-preserving if $i = 0$ and orientation-reversing if $i = 1$, we must have $i = j$, and so we can cancel r from both sides if necessary, to obtain the equality $t_a\rho_\theta = t_b\rho_\eta$. Multiplying both sides on the left by t_{-b} and on the right by $\rho_{-\theta}$, we find $t_{a-b} = \rho_{\eta-\theta}$. But a translation is not a rotation unless both are the trivial operations. So $a = b$ and $\theta = \eta$. □

Computation in M can be done with the symbols t_a, ρ_θ, r using rules for composing them which can be calculated from the formulas (2.3). The necessary rules are as follows:

$$\begin{aligned}
&t_a t_b = t_{a+b}, \quad \rho_\theta\rho_\eta = \rho_{\theta+\eta}, \quad rr = 1,\\
&\rho_\theta t_a = t_{a'}\rho_\theta, \quad \text{where } a' = \rho_\theta(a),\\
&r t_a = t_{a'} r, \quad \text{where } a' = r(a),\\
&r\rho_\theta = \rho_{-\theta} r.
\end{aligned} \tag{2.5}$$

Using these rules, we can reduce any product of our generators to one of the two forms (2.4). The form we get is uniquely determined, because there is only one expression of the form (2.4) for a given motion.

Proof of Theorem (2.2). Let m be a rigid motion which preserves orientation but is not a translation. We want to prove that m is a rotation about some point. It is clear that an orientation-preserving motion which fixes a point p in the plane must be a rotation about p. So we must show that every orientation-preserving motion m which is not a translation fixes some point. We write $m = t_a\rho_\theta$ as in (2.4). By assumption, $\theta \neq 0$. One can use the geometric picture in Figure (2.6) to find the fixed point. In it, ℓ is the line through the origin and perpendicular to a, and the sector with angle θ is situated so as to be bisected by ℓ. The point p is determined by inserting the vector a into the sector, as shown. To check that m fixes p, remember that the operation ρ_θ is the one which is made first, and is followed by t_a.



(2.6) **Figure.** The fixed point of an orientation-preserving motion.

Another way to find the fixed point is by solving the equation $x = t_a\rho_\theta(x)$ algebraically for x. By definition of a translation, $t_a(\rho_\theta(x)) = \rho_\theta(x) + a$. So the equation we need to solve is

$$x - \rho_\theta(x) = a \quad \text{or}$$

$$\begin{bmatrix} 1-\cos\theta & \sin\theta \\ -\sin\theta & 1-\cos\theta \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}. \tag{2.7}$$

Note that $\det(1 - \rho_\theta) = 2 - 2\cos\theta$. The determinant is not zero if $\theta \neq 0$, so there is a unique solution for x.

(2.8) **Corollary.** The motion $m = t_a\rho_\theta$ is the rotation through the angle θ about its fixed point.

Proof. As we just saw, the fixed point of m is the one which satisfies the relation $p = \rho_\theta(p) + a$. Then for any x,

$$m(p + x) = t_a\rho_\theta(p + x) = \rho_\theta(p + x) + a = \rho_\theta(p) + \rho_\theta(x) + a = p + \rho_\theta(x).$$

Thus m sends $p + x$ to $p + \rho_\theta(x)$. So it is the rotation about p through the angle θ, as required. □

Next, we will show that any orientation-reversing motion $m = t_a\rho_\theta r$ is a glide reflection or a reflection (which we may consider to be a glide reflection having glide vector zero). We do this by finding a line ℓ which is sent to itself by m, and so that the motion of m on ℓ is a translation. It is clear geometrically that an orientation-reversing motion which acts in this way on a line is a glide reflection.

The geometry is more complicated here, so we will reduce the problem in two steps. First, the motion $\rho_\theta r = r'$ is a reflection about a line. The line is the one which intersects the x_1-axis at an angle of $\frac{1}{2}\theta$ at the origin. This is not hard to see, geometrically or algebraically. So our motion m is the product of the translation t_a and the reflection r'. We may as well rotate coordinates so that the x_1-axis becomes

the line of reflection of r'. Then r' becomes our standard reflection r, and the translation t_a remains a translation, though the coordinates of the vector a will have changed. In this new coordinate system, the motion is written as $m = t_a r$, and it acts as

$$m\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + a_1 \\ -x_2 + a_2 \end{bmatrix}.$$

This motion sends the line $x_2 = \frac{1}{2}a_2$ to itself, by the translation $(x_1, \frac{1}{2}a_2)^t \rightsquigarrow (x_1 + a_1, \frac{1}{2}a_2)^t$, and so m is a glide along this line. □

There are two important subgroups of M for which we must introduce notation:

(2.9)

T, the group of translations.

O, the group of orthogonal operators.

The group **O** consists of the motions leaving the origin fixed. It contains the rotations about the origin and reflections about lines through the origin.

Notice that with our choice of coordinates we get a bijective correspondence

(2.10)
$$\mathbb{R}^2 \longrightarrow T$$
$$a \rightsquigarrow t_a.$$

This is an isomorphism of the additive group $(\mathbb{R}^2)^+$ with the subgroup T, because $t_a t_b = t_{a+b}$.

The elements of **O** are linear operators. Again making use of our choice of coordinates, we can associate an element $m \in \mathbf{O}$ to its matrix. Doing so, we obtain an isomorphism

$$O_2 \xrightarrow{\sim} \mathbf{O}$$

from the group O_2 of orthogonal 2×2 matrices to **O** [see Chapter 4 (5.16)].

We can also consider the subgroup of M of motions fixing a point of the plane other than the origin. This subgroup is related to **O** as follows:

(2.11) **Proposition.**

(a) Let p be a point of the plane. Let ρ_θ' denote rotation through the angle θ about p, and let r' denote reflection about the line through p and parallel to the x-axis. Then $\rho_\theta' = t_p \rho_\theta t_p^{-1}$ and $r' = t_p r t_p^{-1}$.

(b) The subgroup of M of motions fixing p is the conjugate subgroup

$$\mathbf{O}' = t_p \mathbf{O} t_p^{-1}.$$

Proof. We can obtain the rotation ρ_θ' in this way: First translate p to the origin, next rotate the plane about the origin through the angle θ, and finally translate the origin back to p:

$$\rho_\theta' = t_p \rho_\theta t_{-p} = t_p \rho_\theta t_p^{-1}.$$

The reflection r' can be obtained in the same way from r:

$$r' = t_p r t_{-p} = t_p r t_p^{-1}.$$

This proves (a). Since every motion fixing p has the form ρ_θ' or $\rho_\theta' r'$ [see the proof of (2.4)], (b) follows from (a). □

There is an important homomorphism φ from M to $\mathbf{O}$ whose kernel is T, which is obtained by dropping the translation from the products (2.4):

(2.12)
$$\begin{aligned} M &\xrightarrow{\varphi} \mathbf{O} \\ t_a\rho_\theta &\rightsquigarrow \rho_\theta \\ t_a\rho_\theta r &\rightsquigarrow \rho_\theta r. \end{aligned}$$

This may look too naive to be a good definition, but formulas (2.5) show that φ is a homomorphism: $(t_a\rho_\theta)(t_b\rho_\eta) = t_a t_{b'} \rho_\theta \rho_\eta = t_{a+b'}\rho_{\theta+\eta}$, hence $\varphi(t_a\rho_\theta t_b\rho_\eta) = \rho_{\theta+\eta}$, and so on. Since T is the kernel of a homomorphism, it is a normal subgroup of M.

Note that we can not define a homomorphism from M to T in this way.

(2.13) **Proposition.** Let p be any point of the plane, and let ρ_θ' denote rotation through the angle θ about p. Then $\varphi(\rho_\theta') = \rho_\theta$. Similarly, if r' is reflection about the line through p and parallel to the x-axis, then $\varphi(r') = r$.

This follows from (2.11a), because t_p is in the kernel of φ. The proposition can also be expressed as follows:

(2.14) *The homomorphism φ does not depend on the choice of origin.* □

3. FINITE GROUPS OF MOTIONS

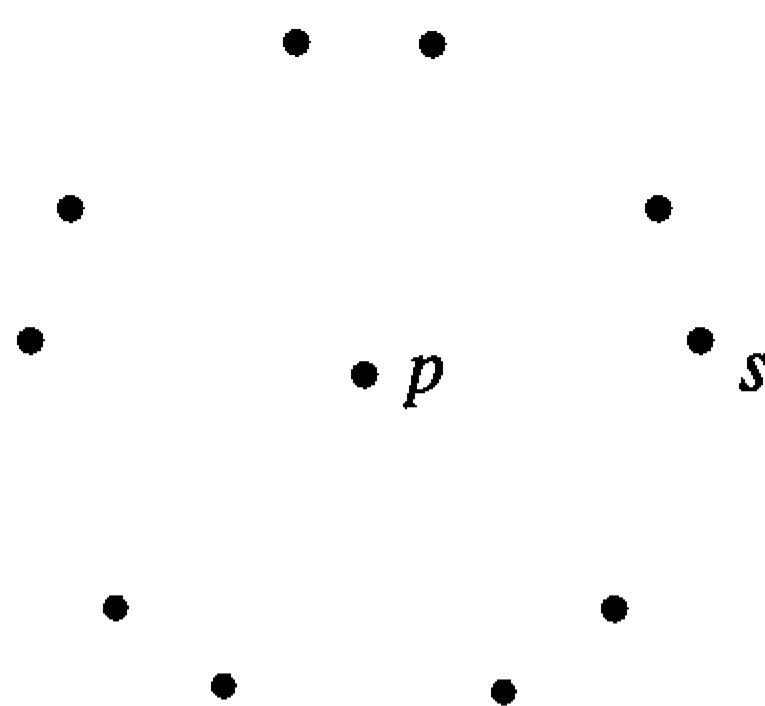
In this section we investigate the possible finite groups of symmetry of figures such as (1.1) and (1.2). So we are led to the study of finite subgroups G of the group M of rigid motions of the plane.

The key observation which allows us to describe all finite subgroups is the following theorem.

(3.1) **Theorem.** *Fixed Point Theorem:* Let G be a finite subgroup of the group of motions M. There is a point p in the plane which is left fixed by every element of G, that is, there is a point p such that $g(p) = p$ for all $g \in G$.

It follows, for example, that any subgroup of M which contains rotations about two different points is infinite.

Here is a beautiful geometric proof of the theorem. Let s be any point in the plane, and let S be the set of points which are the images of s under the various motions in G. So each element $s' \in S$ has the form $s' = g(s)$ for some $g \in G$. This set is called the *orbit* of s under the action of G. The element s is in the orbit because the identity element 1 is in G, and $s = 1(s)$. A typical orbit is depicted below, for the case that G is the group of symmetries of a regular pentagon.



Any element of the group G will permute the orbit S. In other words, if $s' \in S$ and $x \in G$, then $x(s') \in S$. For, say that $s' = g(s)$, with $g \in G$. Since G is a group, $xg \in G$. Therefore, by definition, $xg(s) \in S$. Since $xg(s) = x(s')$, this shows that $x(s') \in S$.

We list the elements of S arbitrarily, writing $S = \{s_1, \dots, s_n\}$. The fixed point we are looking for is the *center of gravity* of the orbit, defined as

$$p = \tfrac{1}{n}(s_1 + \cdots + s_n), \tag{3.2}$$

where the right side is computed by vector addition, using an arbitrary coordinate system in the plane. The center of gravity should be considered an *average* of the points $s_1, \dots, s_n$.

(3.3) **Lemma.** Let $S = \{s_1, \dots, s_n\}$ be a finite set of points of the plane, and let p be its center of gravity, defined by (3.2). Let m be a rigid motion, and let $m(s_i) = s_i'$ and $m(p) = p'$. Then $p' = \frac{1}{n}(s_1' + \cdots + s_n')$. In other words, rigid motions carry centers of gravity to centers of gravity.

Proof. This is clear by physical reasoning. It can also be shown by calculation. To do so, it suffices to treat separately the cases $m = t_a$, $m = \rho_\theta$, and $m = r$, since any motion is obtained from these by composition.

Case 1: $m = t_a$. Then $p' = p + a$ and $s_i' = s_i + a$. It is true that

$$p + a = \tfrac{1}{n}((s_1 + a) + \cdots + (s_n + a)).$$

Case 2: $m = \rho_\theta$ or r. Then m is a linear operator. Therefore

$$p' = m(\tfrac{1}{n}(s_1 + \cdots + s_n)) = \tfrac{1}{n}(m(s_1) + \cdots + m(s_n)) = \tfrac{1}{n}(s_1' + \cdots + s_n'). \qquad \square$$

The center of gravity of our set S is a fixed point for the action of G. For, any element g_i of G permutes the orbit $\{s_1, \ldots, s_n\}$, so Lemma (3.3) shows that it sends the center of gravity to itself. This completes the proof of the theorem. □

Now let G be a finite subgroup of M. Theorem (3.1) tells us that there is a point fixed by every element of G, and we may adjust coordinates so that this point is the origin. Then G will be a subgroup of $\mathbf{O}$. So to describe the finite subgroups G of M, we need only describe the finite subgroups of $\mathbf{O}$ (or, since $\mathbf{O}$ is isomorphic to the group of orthogonal 2×2 matrices, the finite subgroups of the orthogonal group O_2). These subgroups are described in the following theorem.

(3.4) **Theorem.** Let G be a finite subgroup of the group $\mathbf{O}$ of rigid motions which fix the origin. Then G is one of the following groups:

(a) $G = C_n$: the *cyclic group* of order n, generated by the rotation ρ_θ, where $\theta = 2\pi/n$.
(b) $G = D_n$: the *dihedral group* of order $2n$, generated by two elements—the rotation ρ_θ, where $\theta = 2\pi/n$, and a reflection r' about a line through the origin.

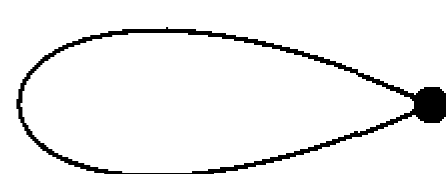
The proof of this theorem is at the end of the section.

The group D_n depends on the line of reflection, but of course we may choose coordinates so that it becomes the x-axis, and then r' becomes our standard reflection r. If G were given as a finite subgroup of M, we would also need to shift the origin to the fixed point in order to apply Theorem (3.4). So our end result about finite groups of motions is the following corollary:

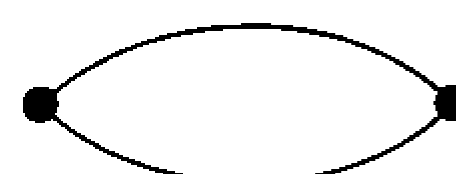
(3.5) **Corollary.** Let G be a finite subgroup of the group of motions M. If coordinates are introduced suitably, then G becomes one of the groups C_n or D_n, where C_n is generated by ρ_θ, $\theta = 2\pi/n$, and D_n is generated by ρ_θ and r. □

When $n \geq 3$, the dihedral group D_n is the group of symmetries of a regular n-sided polygon. This is easy to see, and in fact it follows from the theorem. For a regular n-gon has a group of symmetries which contains the rotation by $2\pi/n$ about its center. It also contains some reflections. Theorem (3.4) tells us that it is D_n.

The dihedral groups D_1, D_2 are too small to be symmetry groups of an n-gon in the usual sense. D_1 is the group $\{1, r\}$ of two elements. So it is a cyclic group, as is C_2. But the nontrivial element of D_1 is a reflection, while in C_2 it is rotation through the angle π. The group D_2 contains the four elements $\{1, \rho, r, \rho r\}$, where $\rho = \rho_\pi$. It is isomorphic to the Klein four group. If we like, we can think of D_1 and D_2 as groups of symmetry of the 1-gon and 2-gon:



1-gon.



2-gon.

The dihedral groups are important examples, and it will be useful to have a complete set of defining relations for them. They can be read off from the list of defining relations for M (2.5). Let us denote the rotation ρ_θ ($\theta = 2\pi/n$) by x, and the reflection r by y.

(3.6) **Proposition.** The dihedral group D_n is generated by two elements x, y which satisfy the relations

$$x^n = 1\ , \quad y^2 = 1\ , \quad yx = x^{-1}y.$$

The elements of D_n are

$$\{1, x, x^2, \ldots, x^{n-1}\ ; y, xy, x^2y, \ldots, x^{n-1}y\} = \{x^iy^j \mid 0 \le i < n, \quad 0 \le j < 2\}\ .$$

Proof. The elements $x = \rho_\theta$ and $y = r$ generate D_n by definition of the group. The relations $y^2 = 1$ and $yx = x^{-1}y$ are included in the list of relations (2.5) for M: They are $rr = 1$ and $r\rho_\theta = \rho_{-\theta}r$. The relation $x^n = 1$ follows from the fact that $\theta = 2\pi/n$, which also shows that the elements $1, x, \ldots, x^{n-1}$ are distinct. It follows that the elements $y, xy, x^2y, \ldots, x^{n-1}y$ are also distinct and, since they are reflections while the powers of x are rotations, that there is no repetition in the list of elements. Finally, the relations can be used to reduce any product of x, y, x^{-1}, y^{-1} to the form x^iy^j , with $0 \le i < n$, $0 \le j < 2$. Therefore the list contains all elements of the group generated by x,y , and since these elements generate D_n the list is complete. □

Using the first two relations (3.6), the third relation can be written in various ways. It is equivalent to

(3.7) $$yx = x^{n-1}y \text{ and also to } xyxy = 1.$$

Note that when $n = 3$, the relations are the same as for the symmetric group S_3 [Chapter 2(1.18)].

(3.8) **Corollary.** The dihedral group D_3 and the symmetric group S_3 are isomorphic. □

For $n > 3$, the dihedral and symmetric groups are certainly not isomorphic, because D_n has order $2n$, while S_n has order $n!$.

Proof of Theorem (3.4). Let G be a finite subgroup of **O**. We need to remember that the elements of **O** are the rotations ρ_θ and the reflections $\rho_\theta r$.

Case 1: All elements of G are rotations. We must prove that G is cyclic in this case. The proof is similar to the determination of the subgroups of the additive group $\mathbb{Z}^+$ of integers [Chapter 2 (2.3)]. If $G = \{1\}$, then $G = C_1$. Otherwise G contains a nontrivial rotation ρ_θ. Let θ be the smallest positive angle of rotation among the elements of G. Then G is generated by ρ_θ. For let ρ_α be any element of G, where the angle of rotation α is represented as usual by a real number. Let $n\theta$ be the greatest integer multiple of θ which is less than α, so that $\alpha = n\theta + \beta$, with $0 \le \beta < \theta$. Since G is a group and since ρ_α and ρ_θ are in G, the product $\rho_\beta = \rho_\alpha\rho_{-n\theta}$ is also in

G. But by assumption θ is the smallest positive angle of rotation in G. Therefore $\beta = 0$ and $\alpha = n\theta$. This shows that G is cyclic. Let $n\theta$ be the smallest multiple of θ which is $\geq 2\pi$, so that $2\pi \leq n\theta < 2\pi + \theta$. Since θ is the smallest positive angle of rotation in G, $n\theta = 2\pi$. Thus $\theta = 2\pi/n$ for some integer n.

Case 2: G contains a reflection. Adjusting coordinates as necessary, we may assume that our standard reflection r is in G. Let H denote the subgroup of rotations in G. We can apply what has been proved in Case 1 to the group H, to conclude that it is a cyclic group: $H = C_n$. Then the $2n$ products $\rho_\theta{}^i$, $\rho_\theta{}^i r$, $0 \leq i \leq n-1$, are in G, and so G contains the dihedral group D_n. We must show that $G = D_n$. Now if an element g of G is a rotation, then $g \in H$ by definition of H; hence g is one of the elements of D_n. If g is a reflection, we can write it in the form $\rho_\alpha r$ for some rotation ρ_α (2.8). Since r is in G, so is the product $\rho_\alpha r r = \rho_\alpha$. Therefore ρ_α is a power of ρ_θ, and g is in D_n too. So $G = D_n$. This completes the proof of the theorem. □

4. DISCRETE GROUPS OF MOTIONS

In this section we will discuss the symmetry groups of unbounded figures such as wallpaper patterns. Our first task is to describe a substitute for the condition that the group is finite—one which includes the groups of symmetry of interesting unbounded figures. Now one property which the patterns illustrated in the text have is that they do not admit arbitrarily small translations or rotations. Very special figures such as a line have arbitrarily small translational symmetries, and a circle, for example, has arbitrarily small rotational symmetries. It turns out that if such figures are ruled out, then the groups of symmetry can be classified.

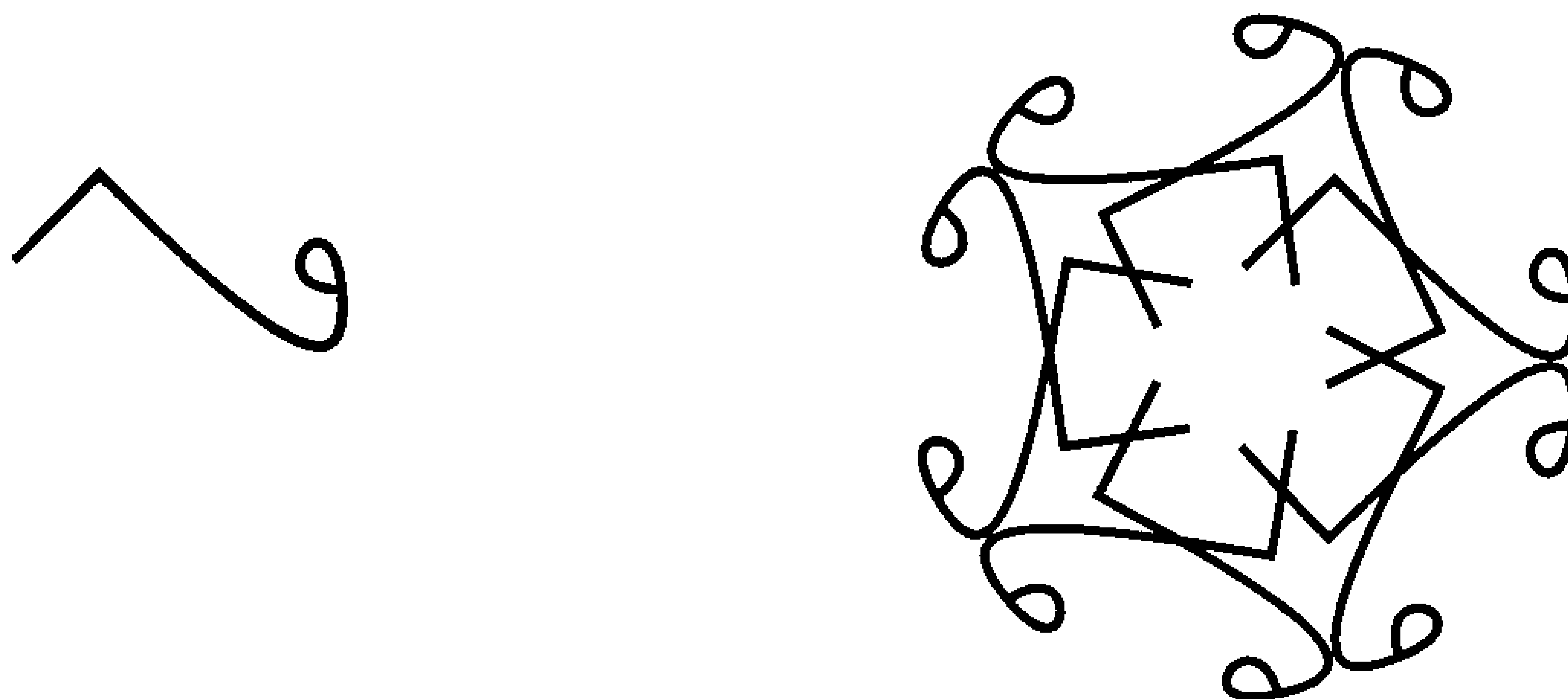
(4.1) **Definition.** A subgroup G of the group of motions M is called *discrete* if it does not contain arbitrarily small translations or rotations. More precisely, G is discrete if there is some real number $\epsilon > 0$ so that

(i) if t_a is a translation in G by a nonzero vector a, then the length of a is at least ϵ: $|a| \geq \epsilon$;
(ii) if ρ is a rotation in G about some point through a nonzero angle θ, then the angle θ is at least ϵ: $|\theta| \geq \epsilon$.

Since the translations and rotations are all the orientation-preserving motions (2.1), this condition applies to all orientation-preserving elements of G. We do not impose a condition on the reflections and glides. The one we might ask for follows automatically from the condition imposed on orientation-preserving motions.

The kaleidoscope principle can be used to show that every discrete group of motions is the group of symmetries of a plane figure. We are not going to give precise reasoning to show this, but the method can be made into a proof. Start with a sufficiently random figure R in the plane. We require in particular that R shall not have any symmetries except for the identity. So every element g of our group will

move R to a different position, call it gR. The required figure F is the union of all the figures gR. An element x of G sends gR to xgR, which is also a part of F, and hence it sends F to itself. If R is sufficiently random, G will be its group of symmetries. As we know from the kaleidoscope, the figure F is often very attractive. Here is the result of applying this procedure in the case that G is the dihedral group of symmetries of a regular pentagon:



Of course many figures have the same group or have similar groups of symmetry. But nevertheless it is interesting and instructive to classify figures according to their groups of symmetry. We are going to discuss a rough classification of the groups, which will be refined in the exercises.

The two main tools for studying a discrete group G are its translation group and its point group. The *translation group* of G is the set of vectors a such that $t_a \in G$. Since $t_a t_b = t_{a+b}$ and $t_{-a} = t_a{}^{-1}$, this is a subgroup of the additive group of vectors, which we will denote by L_G. Using our choice of coordinates, we identify the space of vectors with $\mathbb{R}^2$. Then

(4.2) $$L_G = \{a \in \mathbb{R}^2 \mid t_a \in G\}.$$

This group is isomorphic to the subgroup $T \cap G$ of translations in G, by the isomorphism (2.10): $a \rightsquigarrow t_a$. Since it is a subgroup of G, $T \cap G$ is discrete: A subgroup of a discrete group is discrete. If we translate this condition over to L_G, we find

(4.3) *L_G contains no vector of length $< \epsilon$, except for the zero vector.*

A subgroup L of $\mathbb{R}^{n+}$ which satisfies condition (4.3) for some $\epsilon > 0$ is called a *discrete* subgroup of $\mathbb{R}^n$. Here the adjective *discrete* means that the elements of L are separated by a fixed distance:

(4.4) *The distance between any two vectors $a, b \in L$ is at least ϵ, if $a \neq b$.*

For the distance is the length of $b - a$, and $b - a \in L$ because L is a subgroup.

(4.5) **Proposition.** Every discrete subgroup L of $\mathbb{R}^2$ has one of these forms:

(a) $L = \{0\}$.

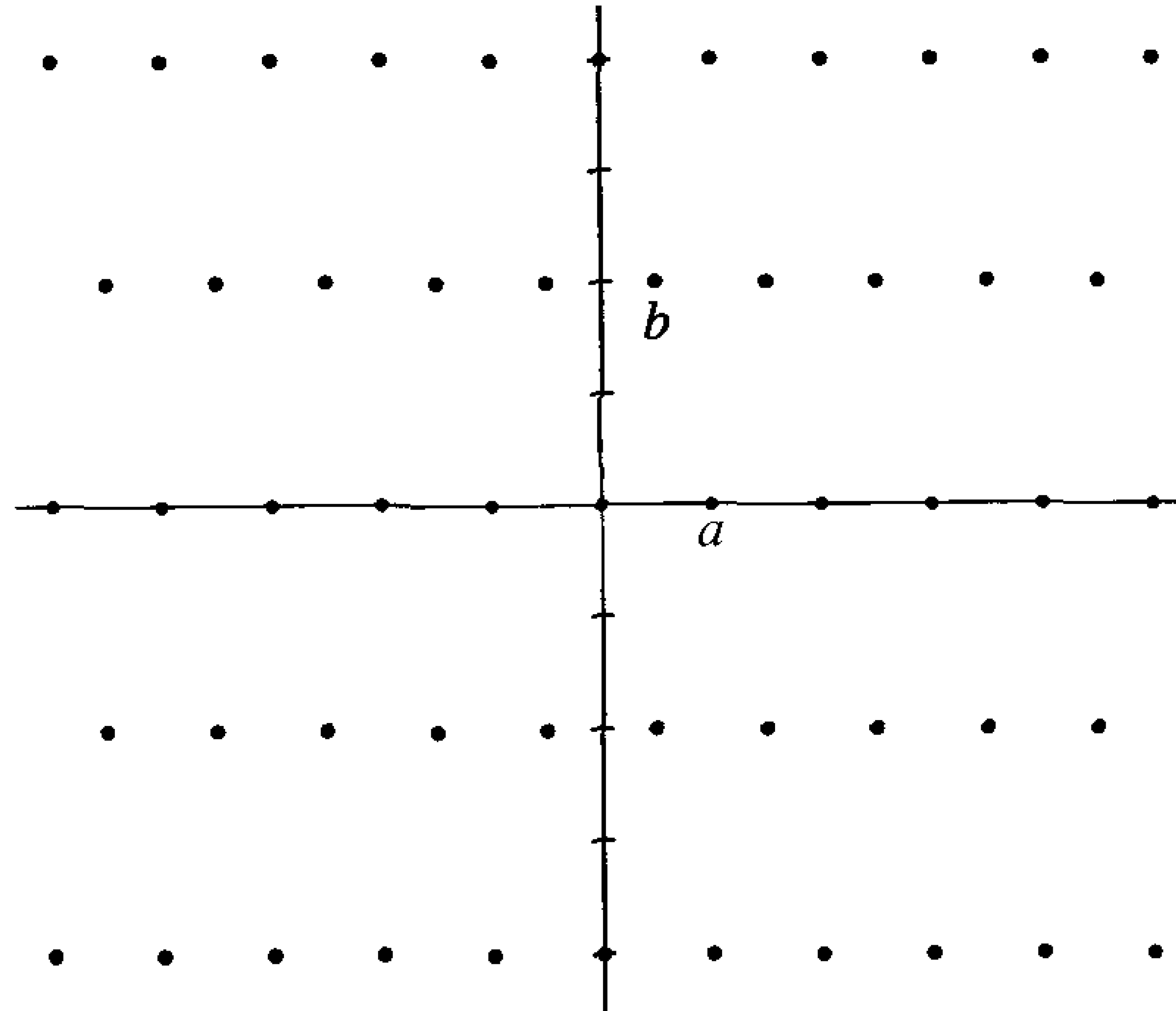
(b) L is generated as an additive group by one nonzero vector a:

$$L = \{ma \mid m \in \mathbb{Z}\}.$$

(c) L is generated by two linearly independent vectors a, b:

$$L = \{ma + nb \mid m,n \in \mathbb{Z}\}.$$

Groups of the third type are called *plane lattices,* and the generating set (a, b) is called a *lattice basis*.



(4.6) **Figure.** A lattice in $\mathbb{R}^2$.

We defer the proof of Proposition (4.5) and turn to the second tool for studying a discrete group of motions—its point group. Recall that there is a homomorphism (2.13) $\varphi: M \longrightarrow \mathbf{O}$, whose kernel is T. If we restrict this homomorphism to G, we obtain a homomorphism

(4.7) $$\varphi|_G: G \longrightarrow \mathbf{O}.$$

Its kernel is $T \cap G$ (which is a subgroup isomorphic to the translation group L_G). The *point group* $\overline{G}$ is the image of G in $\mathbf{O}$. Thus $\overline{G}$ is a subgroup of $\mathbf{O}$.

By definition, a rotation ρ_θ is in $\overline{G}$ if G contains some element of the form $t_a\rho_\theta$. And we have seen (2.8) that $t_a\rho_\theta$ is a rotation through the angle θ about some point in the plane. So the inverse image of an element $\rho_\theta \in \overline{G}$ consists of all of the elements of G which are rotations through the angle θ about some point.

Similarly, let ℓ denote the line of reflection of $\rho_\theta r$. As we have noted before, its angle with the x-axis is $\frac{1}{2}\theta$. The point group $\overline{G}$ contains $\rho_\theta r$ if there is some element $t_a\rho_\theta r$ in G, and $t_a\rho_\theta r$ is a reflection or a glide reflection along a line parallel to ℓ. So the inverse image of $\rho_\theta r$ consists of all elements of G which are reflections and glides along lines parallel to ℓ.

Since G contains no small rotations, the same is true of its point group $\overline{G}$. So $\overline{G}$ is discrete too—it is a discrete subgroup of $\mathbf{O}$.

(4.8) **Proposition.** A discrete subgroup of $\mathbf{O}$ is a finite group.

We leave the proof of this proposition as an exercise. □

Combining Proposition (4.8) with Theorem (3.4), we find the following:

(4.9) **Corollary.** The point group $\overline{G}$ of a discrete group G is cyclic or dihedral. □

Here is the key observation which relates the point group to the translation group:

(4.10) **Proposition.** Let G be a discrete subgroup of M, with translation group $L = L_G$ and point group $\overline{G}$. The elements of $\overline{G}$ carry the group L to itself. In other words, if $\overline{g} \in \overline{G}$ and $a \in L$, then $\overline{g}(a) \in L$.

We may restate this proposition by saying that $\overline{G}$ is contained in the group of symmetries of L, when L is regarded as a set of points in the plane $\mathbb{R}^2$. However, it is important to note that the original group G need not operate on L.

Proof. To say that $a \in L$ means that $t_a \in G$. So we have to show that if $t_a \in G$ and $\overline{g} \in \overline{G}$, then $t_{\overline{g}(a)} \in G$. Now by definition of the point group, $\overline{g}$ is the image of some element g of the group G: $\varphi(g) = \overline{g}$. We will prove the proposition by showing that $t_{\overline{g}(a)}$ is the conjugate of t_a by g. We write $g = t_b\rho$ or $t_b\rho r$, where $\rho = \rho_\theta$. Then $\overline{g} = \rho$ or ρr, according to the case. In the first case,

$$gt_ag^{-1} = t_b\rho t_a\rho^{-1}t_{-b} = t_bt_{\rho(a)}\rho\rho^{-1}t_{-b} = t_{\rho(a)},$$

as required. The computation is similar in the other case. □

The following proposition describes the point groups which can arise when the translation group L_G is a lattice.

(4.11) **Proposition.** Let $H \subset \mathbf{O}$ be a finite subgroup of the group of symmetries of a lattice L. Then

(a) Every rotation in H has order 1, 2, 3, 4, or 6.
(b) H is one of the groups C_n, D_n where $n = 1, 2, 3, 4,$ or 6.

This proposition is often referred to as the *Crystallographic Restriction*. Notice that a rotation of order 5 is ruled out by (4.11). There is no wallpaper pattern with fivefold rotational symmetry. (However, there do exist "quasi-periodic" patterns with fivefold symmetry.)

To prove Propositions (4.5) and (4.11), we begin by noting the following simple lemma:

(4.12) **Lemma.** Let L be a discrete subgroup of $\mathbb{R}^2$.

(a) A bounded subset S of $\mathbb{R}^2$ contains only finitely many elements of L.
(b) If $L \neq \{0\}$, then L contains a nonzero vector of minimal length.

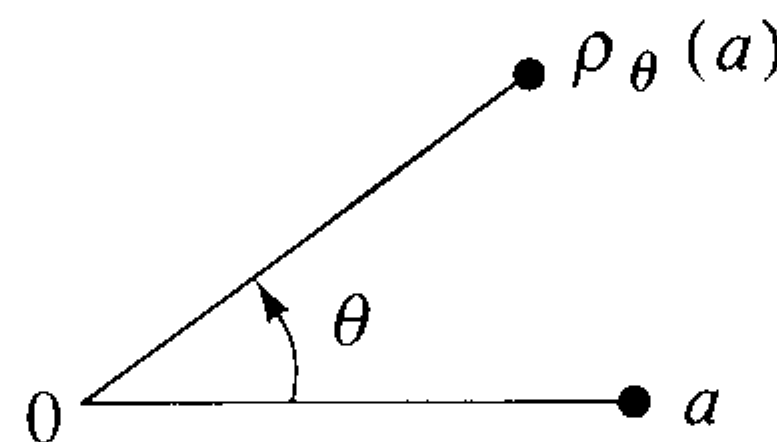
Proof.

(a) Recall that a subset S of $\mathbb{R}^n$ is called bounded if it is contained in some large box, or if the points of S do not have arbitrarily large coordinates. Obviously, if S is bounded, so is $L \cap S$. Now a bounded set which is infinite must contain some elements arbitrarily close to each other—that is, the elements can not be separated by a fixed positive distance ϵ. This is not the case for L, by (4.4). Thus $L \cap S$ is finite.

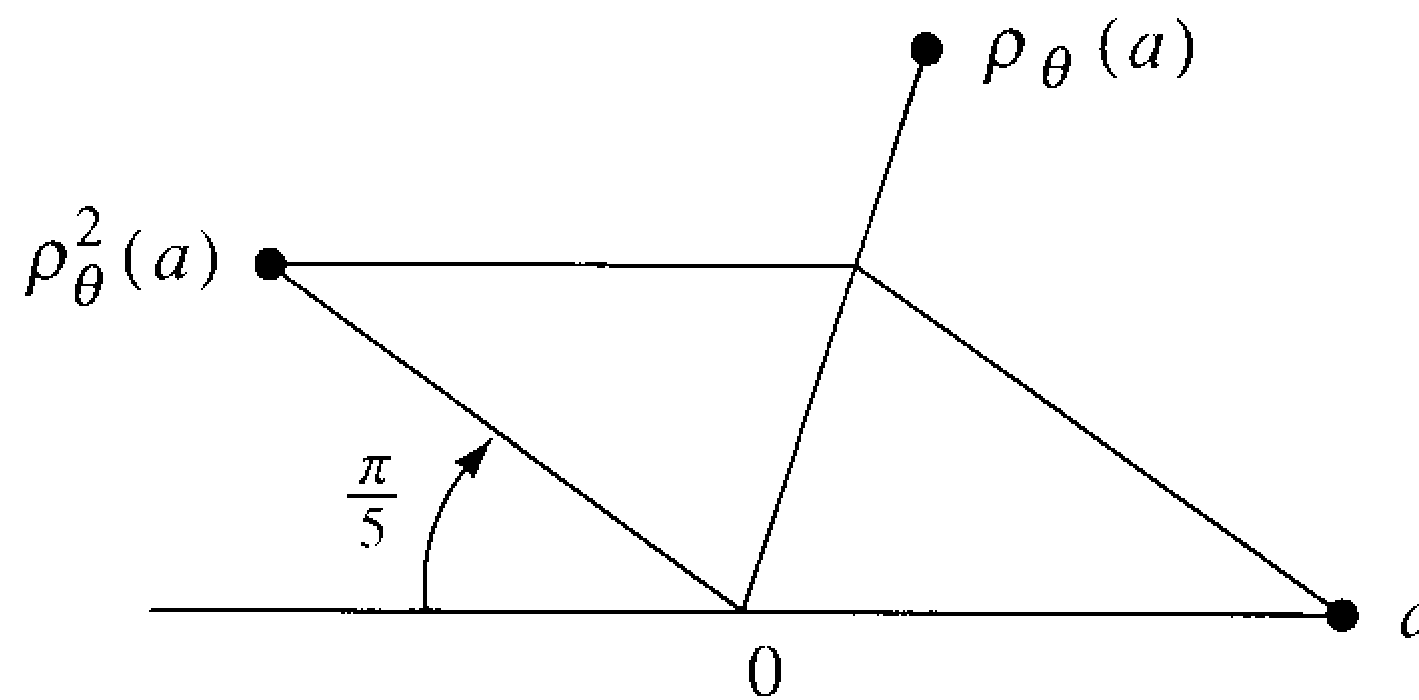
(b) When we say that a nonzero vector a has minimal length, we mean that every nonzero vector $v \in L$ has length at least $|a|$. We don't require the vector a to be uniquely determined. In fact we couldn't require this, because whenever a has minimal length, $-a$ does too.

Assume that $L \neq \{0\}$. To prove that a vector of minimal length exists, we let $b \in L$ be any nonzero vector, and let S be the disc of radius $|b|$ about the origin. This disc is a bounded set, so it contains finitely many elements of L, including b. We search through the nonzero vectors in this finite set to find one having minimal length. It will be the required shortest vector. □

Proof of Proposition (4.11). The second part of the proposition follows from the first, by (3.6). To prove (a), let θ be the smallest nonzero angle of rotation in H, and let a be a nonzero vector in L of minimal length. Then since H operates on L, $\rho_\theta(a)$ is also in L; hence $b = \rho_\theta(a) - a \in L$. Since a has a minimal length, $|b| \geq |a|$. It follows that $\theta \geq 2\pi/6$.



Thus ρ_θ has order ≤ 6. The case that $\theta = 2\pi/5$ is also ruled out, because then the element $b' = \rho_\theta{}^2(a) + a$ is shorter than a:

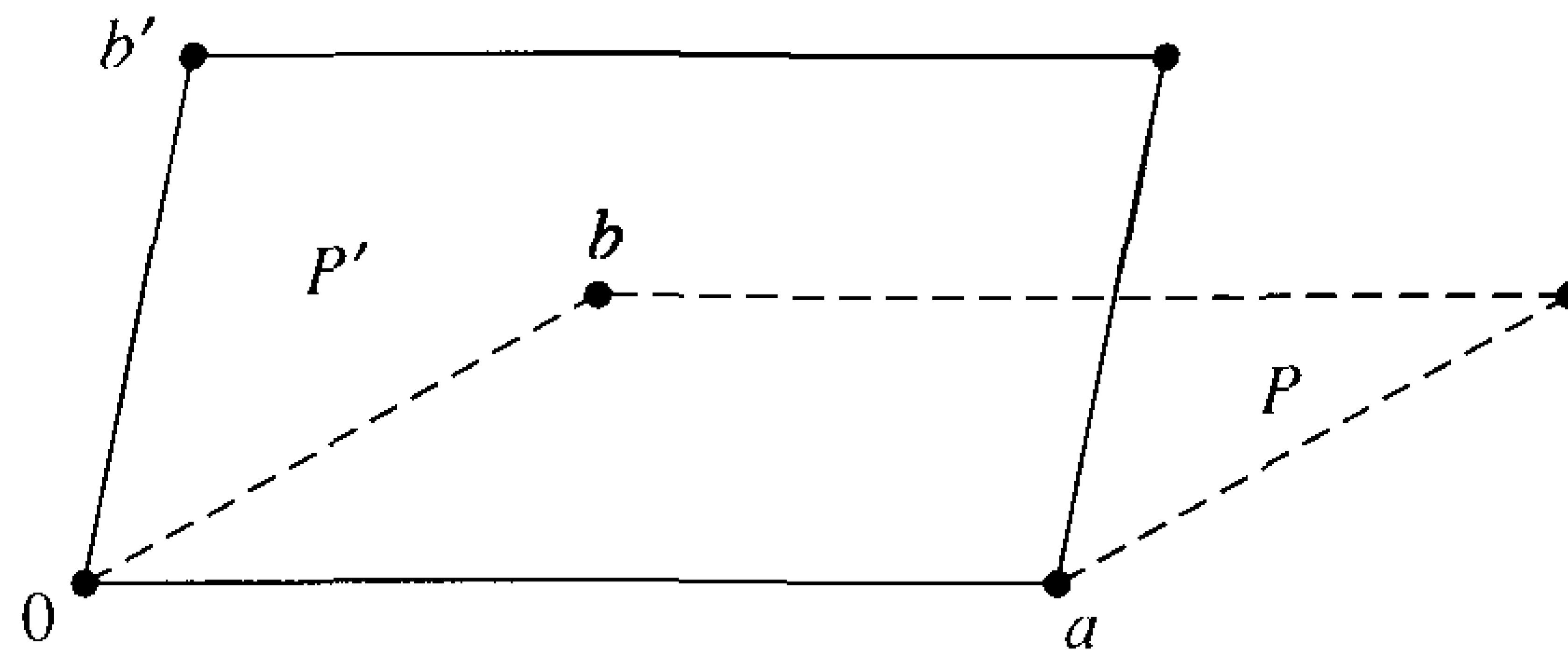


This completes the proof. □

Proof of Proposition (4.5). Let L be a discrete subgroup of $\mathbb{R}^2$. The possibility that $L = \{0\}$ is included in the list. If $L \neq \{0\}$, there is a nonzero vector $a \in L$, and we have two possibilities:

Case 1: All vectors in L lie on one line ℓ through the origin. We repeat an argument used several times before, choosing a nonzero vector $a \in L$ of minimal length. We claim that L is generated by a. Let v be any element of L. Then it is a real multiple $v = ra$ of a, since $L \subset \ell$. Take out the integer part of r, writing $r = n + r_0$, where n is an integer and $0 \leq r_0 < 1$. Then $v - na = r_0 a$ has length less than a, and since L is a group this element is in L. Therefore $r_0 = 0$. This shows that v is an integer multiple of a, and hence that it is in the subgroup generated by a, as required.

Case 2: The elements of L do not lie on a line. Then L contains two linearly independent vectors a', b'. We start with an arbitrary pair of independent vectors, and we try to replace them by vectors which will generate the group L. To begin with, we replace a' by a shortest nonzero vector a on the line ℓ which a' spans. When this is done, the discussion of Case 1 shows that the subgroup $\ell \cap L$ is generated by a. Next, consider the parallelogram P' whose vertices are $0, a, b', a + b'$:



(4.13) **Figure.**

Since P' is a bounded set, it contains only finitely many elements of L (4.12). We may search through this finite set and choose a vector b whose distance to the line ℓ is as small as possible, but positive. We replace b' by this vector. Let P be the parallelogram with $0, a, b, a + b$. We note that P contains no points of L except for its vertices. To see this, notice first that any lattice point c in P which is not a vertex must lie on one of the line segments $[b, a + b]$ or $[0, a]$. Otherwise the two points c and $c - a$ would be closer to ℓ than b, and one of these points would lie in P'. Next, the line segment $[0, a]$ is ruled out by the fact that a is a shortest vector on ℓ. Finally, if there were a point c on $[b, a + b]$, then $c - b$ would be an element of L on the segment $[0, a]$. The proof is completed by the following lemma.

(4.14) **Lemma.** Let a, b be linearly independent vectors which are elements of a subgroup L of $\mathbb{R}^2$. Suppose that the parallelogram P which they span contains no element of L other than the vertices $0, a, b, a + b$. Then L is generated by a and b, that is,

$$L = \{ma + nb \mid m, n \in \mathbb{Z}\}.$$

Proof. Let v be an arbitrary element of L. Then since (a,b) is a basis of $\mathbb{R}^2$, v is a linear combination, say $v = ra + sb$, where r,s are real numbers. We take out the integer parts of r, s, writing $r = m + r_0$, $s = n + s_0$, where m, n are integers and $0 \leq r_0, s_0 < 1$. Let $v_0 = r_0 a + s_0 b = v - ma - nb$. Then v_0 lies in the paral-

lelogram P, and $v_0 \in L$. Hence v_0 is one of the vertices, and since $r_0, s_0 < 1$, it must be the origin. Thus $v = ma + nb$. This completes the proof of the lemma and of Proposition (4.5). □

Let L be a lattice in $\mathbb{R}^2$. An element $v \in L$ is called *primitive* if it is not an integer multiple of another vector in L. The preceding proof actually shows the following:

(4.15) **Corollary.** Let L be a lattice, and let v be a primitive element of L. There is an element $w \in L$ so that the set (v, w) is a lattice basis. □

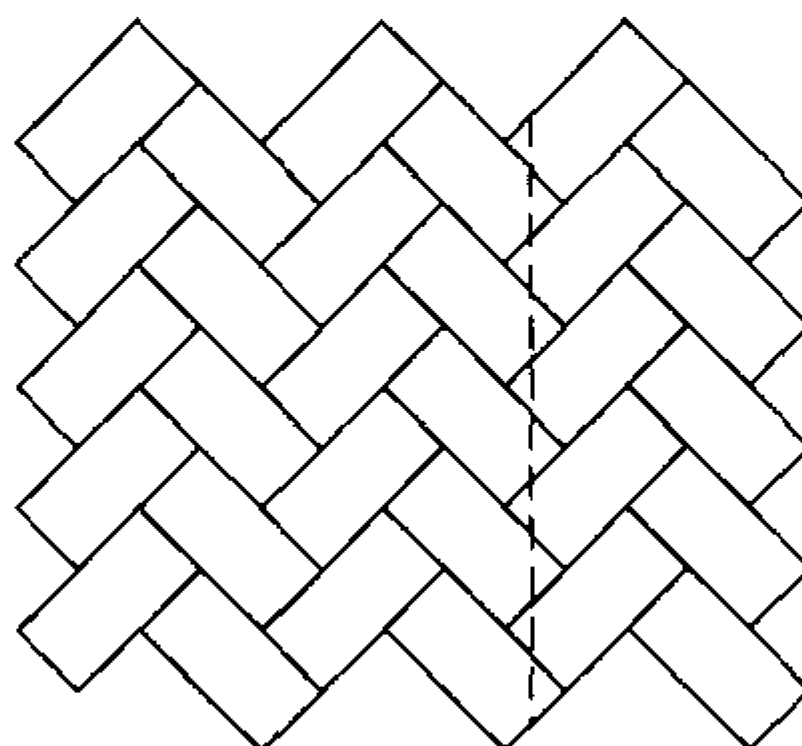
Now let us go back to our discrete group of motions $G \subset M$ and consider the rough classification of G according to the structure of its translation group L_G. If L_G is the trivial group, then the homomorphism from G to its point group is bijective and G is finite. We examined this case in Section 3.

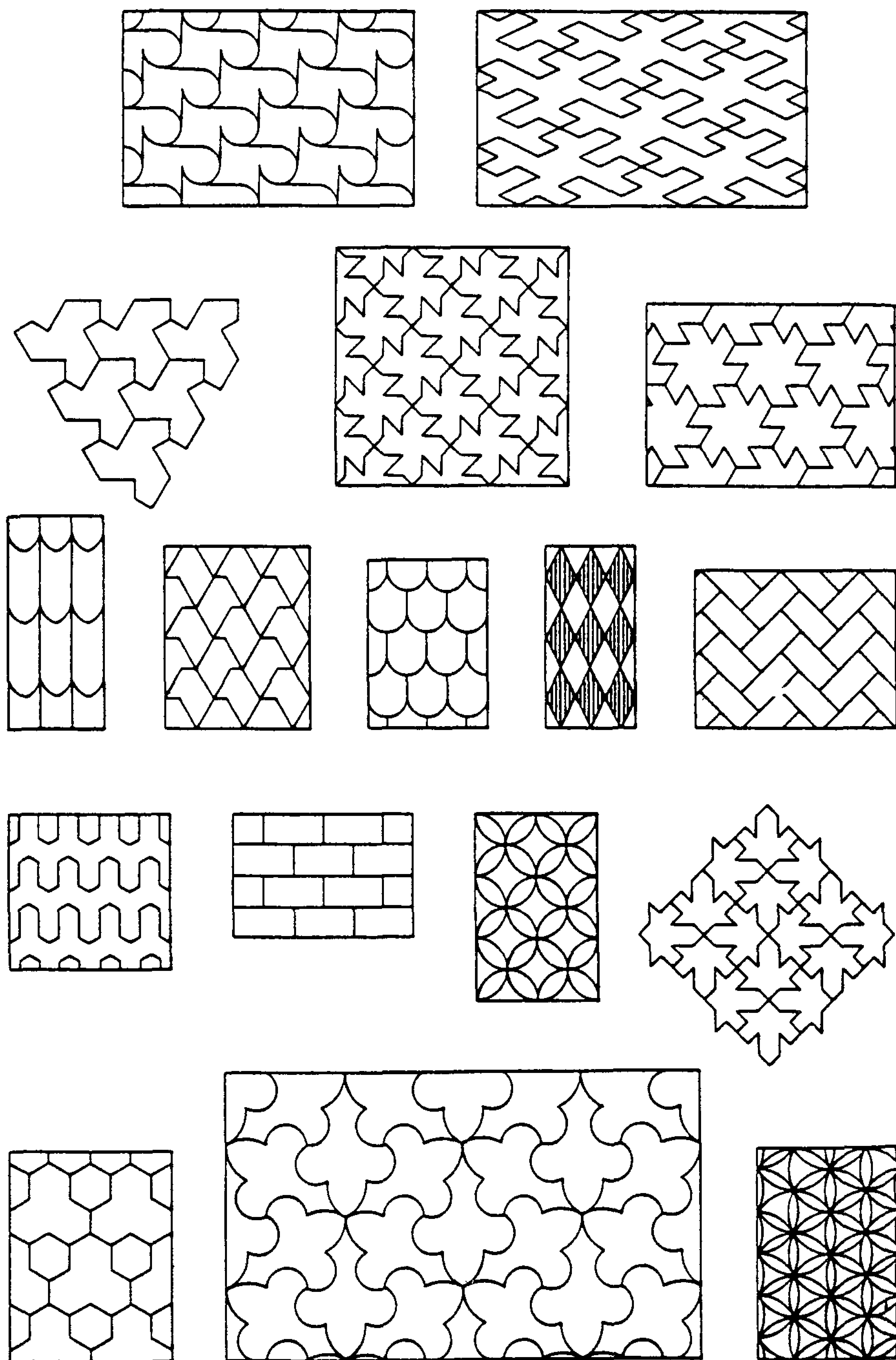
The discrete groups G such that L_G is infinite cyclic are the symmetry groups of frieze patterns such as (1.3). The classification of these groups is left as an exercise.

If L_G is a lattice, then G is called a *two-dimensional crystallographic group*, or a *lattice group*. These groups are the groups of symmetries of wallpaper patterns and of two-dimensional crystals.

The fact that any wallpaper pattern repeats itself in two different directions is reflected in the fact that its group of symmetries will always contain two independent translations, which shows that L_G is a lattice. It may also contain further elements—rotations, reflections, or glides—but the crystallographic restriction limits the possibilities and allows one to classify crystallographic groups into 17 types. The classification takes into account not only the intrinsic structure of the group, but also the type of motion that each group element represents. Representative patterns with the various types of symmetry are illustrated in Figure (4.16).

Proposition (4.11) is useful for determining the point group of a crystallographic group. For example, the brick pattern shown below has a rotational symmetry through the angle π about the centers of the bricks. All of these rotations represent the same element ρ_π of the point group $\overline{G}$. The pattern also has glide symmetry along the dotted line indicated. Therefore the point group $\overline{G}$ contains a reflection. By Proposition (4.11), $\overline{G}$ is a dihedral group. On the other hand, it is easy to see that the only nontrivial rotations in the group G of symmetries are through the angle π. Therefore $\overline{G} = D_2 = \{1, \rho_\pi, r, \rho_\pi r\}$.





(4.16) **Figure.** Sample patterns for the 17 plane crystallographic groups.

The point group $\overline{G}$ and the translation group L_G do not completely characterize the group G. Things are complicated by the fact that a reflection in $\overline{G}$ need not be the image of a reflection in G—it may be represented in G only by glides, as in the brick pattern illustrated above.

As a sample of the methods required to classify the two-dimensional crystallographic groups, we will describe those whose point group contains a rotation ρ through the angle $\pi/2$. According to Proposition (4.11), the point group will be either C_4 or D_4. Since any element of G which represents ρ is also a rotation through $\pi/2$ about some point p, we may choose p to be the origin. Then ρ can be thought of as an element of G too.

(4.17) **Proposition.** Let G be a lattice group whose point group contains a rotation ρ through the angle $\pi/2$. Choose coordinates so that the origin is a point of rotation by $\pi/2$ in G. Let a be a shortest vector in $L = L_G$, let $b = \rho(a)$, and let $c = \frac{1}{2}(a + b)$. Denote by r the reflection about the line spanned by a. Then G is generated by one of the following sets: $\{t_a, \rho\}$, $\{t_a, \rho, r\}$, $\{t_a, \rho, t_c r\}$. Thus there are three such groups.

Proof. We first note that L is a square lattice, generated by a and b. For, a is in L by hypothesis, and Proposition (4.10) asserts that $b = \rho(a)$ is also in L. These two vectors generate a square sublattice L' of L. If $L \neq L'$, then according to Lemma (4.14) there is an element $w \in L$ in the square whose vertices are $0, a, a + b$ and which is not one of the vertices. But any such vector would be at a distance less than $|a|$ from at least one of the vertices v, and the difference $w - v$ would be in L but shorter than a, contrary to the choice of a. Thus $L = L'$, as claimed.

Now the elements t_a and ρ are in G, and $\rho t_a \rho^{-1} = t_b$ (2.5). So the subgroup H of G generated by the set $\{t_a, \rho\}$ contains t_a and t_b. Hence it contains t_w for every $w \in L$. The elements of this group are the products $t_w \rho^i$:

$$H = \{t_w \rho^i \mid w \in L, 0 \leq i \leq 3\}.$$

This is one of our groups. We now consider the possible additional elements which G may contain.

Case 1: Every element of G preserves orientation. In this case, the point group is C_4. Every element of G has the form $m = t_u \rho_\theta$, and if such an element is in G then ρ_θ is in the point group. So $\rho_\theta = \rho^i$ for some i, and $m\rho^{-i} = t_u \in G$ too. Therefore $u \in L$, and $m \in H$. So $G = H$ in this case.

Case 2: G contains an orientation-reversing motion. In this case the point group is D_4, and it contains the reflection about the line spanned by a. We choose coordinates so that this reflection becomes our standard reflection r. Then r will be represented in G by an element of the form $m = t_u r$.

Case 2a: The element u is in L; that is, $t_u \in G$. Then $r \in G$ too, so G contains its point group $\overline{G} = D_4$. If $m' = t_w \rho_\theta$ or if $t_w \rho_\theta r$ is any element of G, then $\rho_\theta r$ is in G

too; hence $t_w \in G$, and $w \in L$. Therefore G is the group generated by the set $\{t_a, \rho, r\}$.

Case 2b: The element u is not in L. This is the hard case.

(4.18) **Lemma.** Let U be the set of vectors u such that $t_u r \in G$. Then

(a) $L + U = U$.
(b) $\rho U = U$.
(c) $U + rU \subset L$.

Proof. If $v \in L$ and $u \in U$, then t_v and $t_u r$ are in G; hence $t_v t_u r = t_{v+u} r \in G$. This shows that $c + v \in U$ and proves (a). Next, suppose that $u \in U$. Then $\rho t_u r \rho = t_{\rho u} \rho r \rho = t_{\rho u} r \in G$. This shows that $\rho u \in U$ and proves (b). Finally, if $u, v \in U$, then $t_u r t_v r = t_{u+rv} \in G$; hence $u + rv \in L$, which proves (c). □

Part (a) of the lemma allows us to choose an element $u \in U$ lying in the square whose vertices are $0, a, b, a + b$ and which is not on the line segments $[a, a + b]$ and $[b, a + b]$. We write u in terms of the basis (a, b), say $u = xa + yb$, where $0 \leq x, y < 1$. Then $u + ru = 2xa$. Since $u + ru \in L$ by (4.18c), the possible values for x are $0, \frac{1}{2}$. Next, $\rho u + a = (1 - y)a + xb$ lies in the square too, and the same reasoning shows that y is 0 or $\frac{1}{2}$. Thus the three possibilities for u are $\frac{1}{2}a$, $\frac{1}{2}b$, and $\frac{1}{2}(a + b) = c$. But if $u = \frac{1}{2}a$, then $\rho u = \frac{1}{2}b$, and $ru = u = \frac{1}{2}a$. So $c = \frac{1}{2}(a + b) \in L$ (4.18b,c). This is impossible because c is shorter than a. Similarly, the case $u = \frac{1}{2}b$ is impossible. So the only remaining case is $u = c$, which means that the group G is generated by $\{t_a, \rho, t_c r\}$. □

5. ABSTRACT SYMMETRY: GROUP OPERATIONS

The concept of symmetry may be applied to things other than geometric figures. For example, complex conjugation $(a + bi) \rightsquigarrow (a - bi)$ may be thought of as a symmetry of the complex numbers. It is compatible with most of the structure of $\mathbb{C}$: If $\overline{\alpha}$ denotes the complex conjugate of α, then $\overline{\alpha + \beta} = \overline{\alpha} + \overline{\beta}$ and $\overline{\alpha\beta} = \overline{\alpha}\,\overline{\beta}$. Being compatible with addition and multiplication, conjugation is called an *automorphism* of the field $\mathbb{C}$. Of course, this symmetry is just the bilateral symmetry of the complex plane about the real axis, but the statement that it is an automorphism refers to its algebraic structure.

Another example of abstract "bilateral" symmetry is given by a cyclic group H of order 3. We saw in Section 3 of Chapter 2 that this group has an automorphism φ, which interchanges the two elements different from the identity.

The set of automorphisms of a group H (or of any other mathematical structure H) forms a group $\operatorname{Aut} H$, the law of composition being composition of maps. Each automorphism should be thought of as a *symmetry* of H, in the sense that it is a permutation of the elements of H which is compatible with the structure of H. But in-

stead of being a geometric figure with a rigid shape, the structure in this case is the group law. The group of automorphisms of the cyclic group of order 3 contains two elements: the identity map and the map φ.

So the words *automorphism* and *symmetry* are more or less synonymous, except that automorphism is used to describe a permutation of a set which preserves some algebraic structure, while symmetry often refers to a permutation which preserves a geometric structure.

These examples are special cases of a more general concept, that of an operation of a group on a set. Suppose we are given a group G and a set S. An *operation* of G on S is a rule for combining elements $g \in G$ and $s \in S$ to get an element gs of S. In other words, it is a law of composition, a map $G \times S \longrightarrow S$, which we generally write as multiplication:

$$g, s \rightsquigarrow gs.$$

This rule is required to satisfy the following axioms:

(5.1)

(a) $1s = s$ for all s (1 is the identity of G).

(b) *Associative law:* $(gg')s = g(g's)$, for all $g, g' \in G$ and $s \in S$.

A set S with an operation of G is often called a *G-set*. This should really be called a *left operation,* because elements of G multiply on the left.

Examples of this concept can be found manywhere. For example, let $G = M$ be the group of all rigid motions of the plane. Then M operates on the set of points of the plane, on the set of lines in the plane, on the set of triangles in the plane, and so on. Or let G be the cyclic group $\{1, r\}$ of order 2, with $r^2 = 1$. Then G operates on the set S of complex numbers, by the rule $r\alpha = \overline{\alpha}$. The fact that the axioms (5.1) hold in a given example is usually clear.

The reason that such a law of composition is called an operation is this: If we fix an element g of G but let $s \in S$ vary, then *left multiplication by g* defines a map from S to itself; let us denote this map by m_g. Thus

(5.2) $$m_g: S \longrightarrow S$$

is defined by

$$m_g(s) = gs.$$

This map describes the way the element g operates on S. Note that m_g is a *permutation* of S; that is, it is bijective. For the axioms show that it has the two-sided inverse

$$m_{g^{-1}} = \textit{multiplication by } g^{-1}:$$

$m_{g^{-1}}(m_g(s)) = g^{-1}(gs) = (g^{-1}g)s = 1s = s$. Interchanging the roles of g and g^{-1} shows that $m_g(m_{g^{-1}}(s)) = s$ too.

The main thing that we can do to study a set S on which a group G operates is to decompose the set into orbits. Let s be an element of S. The *orbit* of s in S is the set

(5.3) $$O_s = \{s' \in S \mid s' \in gs \text{ for some } g \in G\}.$$

It is a subset of S. (The orbit is often written as $Gs = \{gs \mid g \in G\}$, in analogy with the notation for cosets [Chapter 2 (6.1)]. We won't do this because Gs looks too much like the notation for a stabilizer which we are about to introduce.) If we think of elements of G as operating on S by permutations, then O_s is the set of images of s under the various permutations m_g. Thus, if $G = M$ is the group of motions and S is the set of triangles in the plane, the orbit O_Δ of a given triangle Δ is the set of all triangles congruent to Δ. Another example of orbit was introduced when we proved the existence of a fixed point for the operation of a finite group on the plane (3.1).

The orbits for a group action are equivalence classes for the relation

(5.4) $$s \sim s' \text{ if } s' = gs \text{ for some } g \in G.$$

The proof that this is an equivalence relation is easy, so we omit it; we made a similar verification when we introduced cosets in Section 6 of Chapter 2. Being equivalence classes, the orbits partition the set S:

(5.5) *S is a union of disjoint orbits.*

The group G operates on S by operating independently on each orbit. In other words, an element $g \in G$ permutes the elements of each orbit and does not carry elements of one orbit to another orbit. For example, the set of triangles of the plane can be partitioned into congruence classes, the orbits for the action of M. A motion m permutes each congruence class separately. Note that the orbits of an element s and of gs are equal.

If S consists of just one orbit, we say that G operates *transitively* on S. This means that every element of S is carried to every other one by some element of the group. Thus the group of symmetries of Figure (1.7) operates transitively on the set of its legs. The group M of rigid motions of the plane operates transitively on the set of points of the plane, and it operates transitively on the set of lines in the plane. It does not operate transitively on the set of triangles in the plane.

The *stabilizer* of an element $s \in S$ is the subgroup G_s of G of elements leaving s fixed:

(5.6) $$G_s = \{g \in G \mid gs = s\}.$$

It is clear that this is a subgroup. Just as the kernel of a group homomorphism $\varphi: G \longrightarrow G'$ tells us when two elements $x, y \in G$ have the same image, namely, if $x^{-1}y \in \ker \varphi$ [Chapter 2 (5.13)], we can describe when two elements $x, y \in G$ act in the same way on an element $s \in S$ in terms of the stabilizer G_s:

(5.7) $$xs = ys \text{ if and only if } x^{-1}y \in G_s.$$

For $xs = ys$ implies $s = x^{-1}ys$, and conversely.

As an example of a nontrivial stabilizer, consider the action of the group M of rigid motions on the set of points of the plane. The stabilizer of the origin is the subgroup **O** of orthogonal operators.

Or, if S is the set of triangles in the plane and Δ is a particular triangle which happens to be equilateral, then the stabilizer of Δ is its group of symmetries, a subgroup of M isomorphic to D_3 (see (3.4)). Note that when we say that a motion m stabilizes a triangle Δ, we don't mean that m fixes the points of Δ. The only motion which fixes every point of a triangle is the identity. We mean that in permuting the set of triangles, the motion carries Δ to itself. It is important to be clear about this distinction.

6. THE OPERATION ON COSETS

Let H be a subgroup of a group G. We saw in Section 6 of Chapter 2 that the left cosets $aH = \{ah \mid h \in H\}$ form a partition of the group [Chapter 2 (6.3)]. We will call the set of left cosets the *coset space* and will often denote it by G/H, copying this notation from that used for quotient groups when the subgroup is normal.

The fundamental observation to be made is this: Though G/H is not a group unless the subgroup H is normal, nevertheless *G operates on the coset space G/H* in a natural way. The operation is quite obvious: Let g be an element of the group, and let C be a coset. Then gC is defined to be the coset

(6.1) $$gC = \{gc \mid c \in C\}.$$

Thus if $C = aH$, then gC is the coset gaH. It is clear that the axioms (5.1) for an operation are satisfied.

Note that the group G operates transitively on G/H, because G/H is the orbit of the coset $1H = H$. The stabilizer of the coset $1H$ is the subgroup $H \subset G$. Again, note the distinction: Multiplication by an element $h \in H$ does not act trivially on the elements of the coset $1H$, but it sends that coset to itself.

To understand the operation on cosets, you should work carefully through the following example. Let G be the group D_3 of symmetries of an equilateral triangle. As in (3.6), it may be described by generators x,y satisfying the relations $x^3 = 1$, $y^2 = 1$, $yx = x^2y$. Let $H = \{1, y\}$. This is a subgroup of order 2. Its cosets are

(6.2) $$C_1 = H = \{1, y\},\ C_2 = \{x, xy\},\ C_3 = \{x^2, x^2y\},$$

and G operates on $G/H = \{C_1, C_2, C_3\}$. So, as in (5.2), every element g of G determines a permutation m_g of $\{C_1, C_2, C_3\}$. The elements x, y operate as

(6.3) $$m_x: 1 \to 2 \to 3 \to 1 \quad \text{and} \quad m_y: 1 \circlearrowleft \ \ 2 \leftrightarrow 3.$$

In fact, the six elements of G yield all six permutations of three elements, and so the map

$$G \longrightarrow S_3 \approx \mathrm{Perm}(G/H)$$
$$g \rightsquigarrow m_g$$

is an isomorphism. Thus the dihedral group $G = D_3$ is isomorphic to the symmetric group S_3. We already knew this.

The following proposition relates an arbitrary group operation to the operation on cosets:

(6.4) **Proposition.** Let S be a G-set, and let s be an element of S. Let H be the stabilizer of s, and let O_s be the orbit of s. There is a natural bijective map

$$G/H \xrightarrow{\varphi} O_s$$

defined by

$$aH \rightsquigarrow as.$$

This map is compatible with the operations of G in the sense that $\varphi(gC) = g\varphi(C)$ for every coset C and every element $g \in G$.

The proposition tells us that every group operation can be described in terms of the operations on cosets. For example, let $S = \{v_1, v_2, v_3\}$ be the set of vertices of an equilateral triangle, and let G be the group of its symmetries, presented as above. The element y is a reflection which stabilizes one of the vertices of the triangle, say v_1. The stabilizer of this vertex is $H = \{1, y\}$, and its orbit is S. With suitable indexing, the set (6.2) of cosets maps to S by the map $C_i \rightsquigarrow v_i$.

Proof of Proposition (6.4). It is clear the map φ, if it exists, will be compatible with the operation of the group. What is not so clear is that the rule $gH \rightsquigarrow gs$ defines a map at all. Since many symbols gH represent the same coset, we must show that if a and b are group elements and if $aH = bH$, then $as = bs$ too. This is true, because we know that $aH = bH$ if and only if $b = ah$ for some h in H [Chapter 2 (6.5)]. And when $b = ah$, then $bs = ahs = as$, because h fixes s. Next, the orbit of s consists of the elements gs, and φ carries gH to gs. Thus φ maps G/H onto O_s, and φ is surjective. Finally, we show that φ is injective. Suppose aH and bH have the same image: $as = bs$. Then $s = a^{-1}bs$. Since H was defined to be the stabilizer of s, this implies that $a^{-1}b = h \in H$. Thus $b = ah \in aH$, and so $aH = bH$. This completes the proof. □

(6.5) **Proposition.** Let S be a G-set, and let $s \in S$. Let s' be an element in the orbit of s, say $s' = as$. Then

(a) The set of elements g of G such that $gs = s'$ is the left coset

$$aG_s = \{g \in G \mid g = ah \text{ for some } h \in G_s\}.$$

(b) The stabilizer of s' is a *conjugate subgroup* of the stabilizer of s:

$$G_{s'} = aG_s a^{-1} = \{g \in G \mid g = aha^{-1} \text{ for some } h \in G_s\}.$$

We omit the proof. □

As an example, let us recompute the stabilizer of a point p in the plane, for the operation of the group of motions. We have made this computation before, in (2.11b). We have $p = t_p(0)$, and the stabilizer of the origin is the orthogonal group $\mathbf{O}$. Thus by (6.5b),

$$G_p = t_p \mathbf{O} t_p^{-1} = t_p \mathbf{O} t_{-p} = \{m \in M \mid m = t_p \rho_\theta t_{-p} \text{ or } m = t_p \rho_\theta r t_{-p}\}.$$

We know on the other hand that G_p consists of rotations and reflections about the point p. Those are the motions fixing p. So $t_p \mathbf{O} t_p^{-1}$ consists of these elements. This agrees with (2.11).

7. THE COUNTING FORMULA

Let H be a subgroup of G. As we know from Chapter 2 (6.9), all the cosets of H in G have the same number of elements: $|H| = |aH|$. Since G is a union of nonoverlapping cosets and the number of cosets is the index, which we write as $[G:H]$ or $|G/H|$, we have the fundamental formula for the order $|G|$ of the group G (see [Chapter 2 (6.10)]):

(7.1) $$|G| = |H||G/H|.$$

Now let S be a G-set. Then we can combine Proposition (6.4) with (7.1) to get the following:

(7.2) **Proposition.** *Counting Formula*: Let $s \in S$. Then

$$(\text{order of } G) = (\text{order of stabilizer})(\text{order of orbit})$$

$$|G| = |G_s||O_s|.$$

Equivalently, the order of the orbit is equal to the index of the stabilizer:

$$|O_s| = [G : G_s].$$

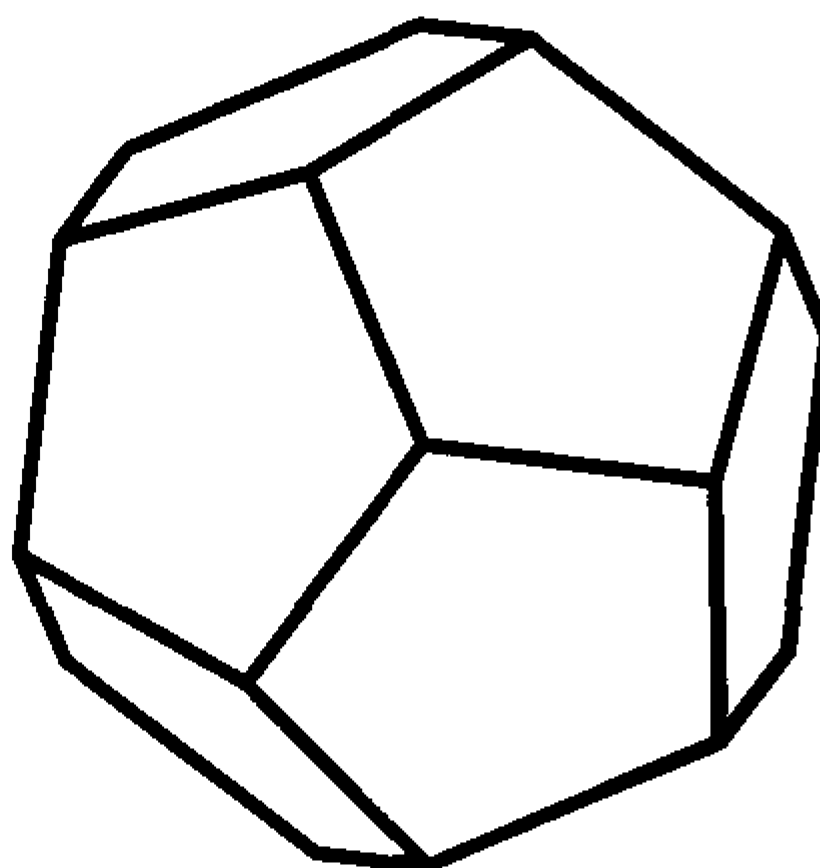
There is one such equation for every $s \in S$. As a consequence, the order of an orbit divides the order of the group.

A more elementary formula uses the partition of S into orbits to count its elements. We label the different orbits which make up S in some way, say as $O_1, \ldots, O_k$. Then

(7.3) $$|S| = |O_1| + |O_2| + \cdots + |O_k|.$$

These simple formulas have a great number of applications.

(7.4) **Example.** Consider the group G of orientation-preserving symmetries of a regular dodecahedron D. It follows from the discussion of Section 8 of Chapter 4 that these symmetries are all rotations. It is tricky to count them without error. Consider the action of G on the set S of the faces of D. The stabilizer of a face s is the group of rotations by multiples of $2\pi/5$ about a perpendicular through the center of s. So the order of G_s is 5. There are 12 faces, and G acts transitively on them. Thus $|G| = 5 \cdot 12 = 60$. Or, G operates transitively on the vertices v of D. There are three rotations, including 1, which fix a vertex, so $|G_v| = 3$. There are 20 vertices; hence $|G| = 3 \cdot 20 = 60$, which checks. There is a similar computation for edges. If e is an edge, then $|G_e| = 2$, so since $60 = 2 \cdot 30$, the dodecahedron has 30 edges.



Following our general principle, we should study restriction of an operation of a group G to a subgroup. Suppose that G operates on a set S, and let H be a subgroup of G. We may restrict the operation, to get an operation of H on S. Doing so leads to more numerical relations.

Clearly, the H-orbit of an element s will be contained in its G-orbit. So we may take a single G-orbit and decompose it into H-orbits. We count the orders of these H-orbits, obtaining another formula. For example, let S be the set of 12 faces of the dodecahedron, and let H be the stabilizer of a particular face s. Then H also fixes the face opposite to s, and so there are two H-orbits of order 1. The remaining faces make up two orbits of order 5. In this case, (7.3) reads as follows.

$$12 = 1 + 1 + 5 + 5.$$

Or let S be the set of faces, and let K be the stabilizer of a vertex. Then K does not fix any face, so every K-orbit has order 3:

$$12 = 3 + 3 + 3 + 3.$$

These relations give us a way of relating several subgroups of a group.

We close the section with a simple application of this procedure to the case that the G-set is the coset space of a subgroup:

(7.5) **Proposition.** Let H and K be subgroups of a group G. Then the index of $H \cap K$ in H is at most equal to the index of K in G:

$$[H : H \cap K] \leq [G : K].$$

Proof. To minimize confusion, let us denote the coset space G/K by S, and the coset $1K$ by s. Thus $|S| = [G : K]$. As we have already remarked, the stabilizer of s is the subgroup K. We now restrict the action of G to the subgroup H and decompose S into H-orbits. The stabilizer of s for this restricted operation is obviously $H \cap K$. We don't know much about the H-orbit O of s except that it is a subset of S. We now apply Proposition (7.2), which tells us that $|O| = [H : H \cap K]$. Therefore $[H : H \cap K] = |O| \leq |S| = [G : K]$, as required. □

8. PERMUTATION REPRESENTATIONS

By its definition, the symmetric group S_n operates on the set $S = \{\mathbf{1},\ldots,\mathbf{n}\}$. A *permutation representation* of a group G is a homomorphism

(8.1)
$$\varphi: G \longrightarrow S_n.$$

Given any such representation, we obtain an operation of G on $S = \{\mathbf{1},\ldots,\mathbf{n}\}$ by letting m_g (5.2) be the permutation $\varphi(g)$. In fact, operations of a group G on $\{\mathbf{1},\ldots,\mathbf{n}\}$ correspond in a bijective way to permutation representations.

More generally, let S be any set, and denote by Perm(S) the group of its permutations. Let G be a group.

(8.2) **Proposition.** There is a bijective correspondence

$$\begin{bmatrix}\text{operations}\\ \text{of } G \text{ on } S\end{bmatrix} \longleftrightarrow \begin{bmatrix}\text{homomorphisms}\\ G \longrightarrow \text{Perm}(S)\end{bmatrix}$$

defined in this way: Given an operation, we define $\varphi: G \longrightarrow \text{Perm}(S)$ by the rule $\varphi(g) = m_g$, where m_g is multiplication by g (5.2).

Let us show that φ is a homomorphism, leaving the rest of the proof of (8.2) as an exercise. We've already noted in Section 5 that m_g is a permutation. So as defined above, $\varphi(g) \in \text{Perm}(S)$. The axiom for a homomorphism is $\varphi(xy) = \varphi(x)\varphi(y)$, or $m_{xy} = m_x m_y$, where multiplication is composition of permutations. So we have to show that $m_{xy}(s) = m_x(m_y(s))$ for every $s \in S$. By Definition (5.2), $m_{xy}(s) = (xy)s$ and $m_x(m_y(s)) = x(ys)$. The associative law (5.1b) for group operations shows that $(xy)s = x(ys)$, as required. □

The isomorphism $D_3 \longrightarrow S_3$ obtained in Section 6 by the action of D_3 on the cosets of H (6.2) is a particular example of a permutation representation. But a homomorphism need not be injective or surjective. If $\varphi: G \longrightarrow \text{Perm}(S)$ happens to be injective, we say that the corresponding operation is *faithful*. So to be faithful, the operation must have the property that $m_g \neq$ identity, unless $g = 1$, or

$$\text{if } gs = s \text{ for every } s \in S, \text{ then } g = 1.$$

The operation of the group of motions M on the set S of equilateral triangles in the plane is faithful, because the identity is the only motion which fixes *all* triangles.

The rest of this section contains a few applications of permutation representations.

(8.3) **Proposition.** The group $GL_2(\mathbb{F}_2)$ of invertible matrices with mod 2 coefficients is isomorphic to the symmetric group S_3.

Proof. Let us denote the field $\mathbb{F}_2$ by F, and the group $GL_2(\mathbb{F}_2)$ by G. We have listed the six elements of G before [Chapter 3 (2.10)]. Let $V = F^2$ be the space of column vectors. This space consists of the following four vectors: $V = \{0, e_1, e_2, e_1 + e_2\}$. The group G operates on V and fixes 0, so it operates on the set of three nonzero vectors, which form one orbit. This gives us a permutation representation $\varphi: G \longrightarrow S_3$. Now the image of e_1 under multiplication by a matrix $P \in G$ is the first column of P, and similarly the image of e_2 is the second column of P. Therefore P can not operate trivially on these two elements unless it is the identity. This shows that the operation of G is faithful, and hence that the map φ is injective. Since both groups have order 6, φ is an isomorphism. □

(8.4) **Proposition.** Let c_g denote conjugation by g, the map $c_g(x) = gxg^{-1}$. The map $f: S_3 \longrightarrow \text{Aut}(S_3)$ from the symmetric group to its group of automorphisms which is defined by the rule $g \rightsquigarrow c_g$ is bijective.

Proof. Let A denote the group of automorphisms of S_3. We know from Chapter 2 (3.4) that c_g is an automorphism. Also, $c_{gh} = c_g c_h$ because $c_{gh}(x) = (gh)x(gh)^{-1} = ghxh^{-1}g^{-1} = c_g(c_h(x))$ for all x. This shows that f is a homomorphism. Now conjugation by g is the identity if and only if g is in the center of the group. The center of S_3 is trivial, so f is injective.

It is to prove surjectivity of f that we look at a permutation representation of A. The group A operates on the set S_3 in the obvious way; namely, if α is an automorphism and $s \in S_3$, then $\alpha s = \alpha(s)$. Elements of S_3 of different orders will be in distinct orbits for this operation. So A operates on the subset of S_3 of elements of order 2. This set contains the three elements $\{y, xy, x^2y\}$. If an automorphism α fixes both xy and y, then it also fixes their product $xyy = x$. Since x and y generate S_3, the only such automorphism is the identity. This shows that the operation of A on $\{y, xy, x^2y\}$ is faithful and that the associated permutation representation $A \longrightarrow \text{Perm}\{y, xy, x^2y\}$ is injective. So the order of A is at most 6. Since f is injective and the order of S_3 is 6, it follows that f is bijective. □

(8.5) **Proposition.** The group of automorphisms of the cyclic group of order p is isomorphic to the multiplicative group $\mathbb{F}_p^{\times}$ of nonzero elements of $\mathbb{F}_p$.

Proof. The method here is to use the additive group $\mathbb{F}_p^{+}$ as the model for a cyclic group of order p. It is generated by the element 1. Let us denote the multiplicative group $\mathbb{F}_p^{\times}$ by G. Then G operates on $\mathbb{F}_p^{+}$ by left multiplication, and this operation defines an injective homomorphism $\varphi: G \longrightarrow \text{Perm}(\mathbb{F}_p)$ to the group of permutations of the set $\mathbb{F}_p$ of p elements.

Next, the group $A = \text{Aut}(\mathbb{F}_p{}^+)$ of automorphisms is a subgroup of $\text{Perm}(\mathbb{F}_p{}^+)$. The distributive law shows that multiplication by an element $a \in \mathbb{F}_p{}^\times$ is an automorphism of $\mathbb{F}_p{}^+$. It is bijective, and $a(x + y) = ax + ay$. Therefore the image of $\varphi: G \longrightarrow \text{Perm}(\mathbb{F}_p{}^+)$ is contained in the subgroup A. Finally, an automorphism of $\mathbb{F}_p{}^+$ is determined by where it sends the generator 1, and the image of 1 can not be zero. Using the operations of G, we can send 1 to any nonzero element. Therefore φ is a surjection from G onto A. Being both injective and surjective, φ is an isomorphism. □

9. FINITE SUBGROUPS OF THE ROTATION GROUP

In this section, we will apply the Counting Formula to classify finite subgroups of the rotation group SO_3, which was defined in Chapter 4 (5.4). As happens with finite groups of motions of the plane, there are rather few finite subgroups of SO_3, and all of them are symmetry groups of familiar figures.

(9.1) **Theorem.** Every finite subgroup G of SO_3 is one of the following:

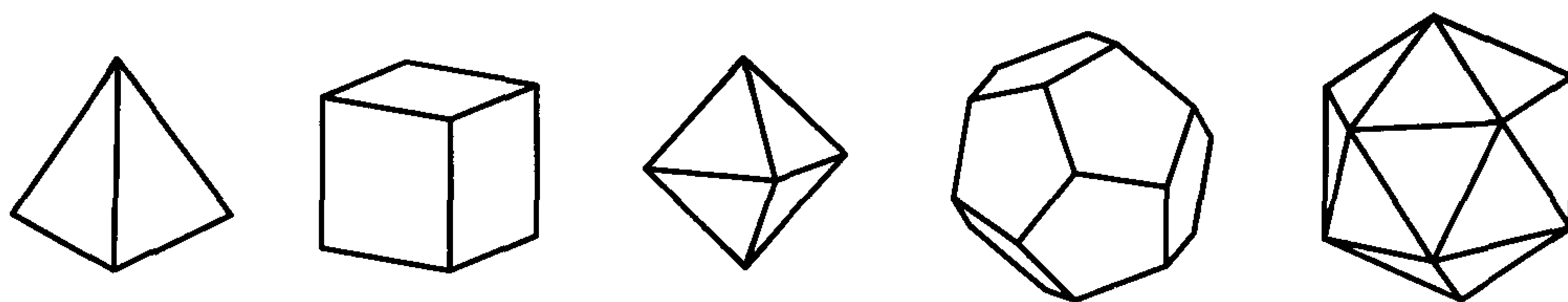
C_k: the *cyclic group* of rotations by multiples of $2\pi/k$ about a line;

D_k: the *dihedral group* (3.4) of symmetries of a regular k–gon;

T: the *tetrahedral group* of twelve rotations carrying a regular tetrahedron to itself;

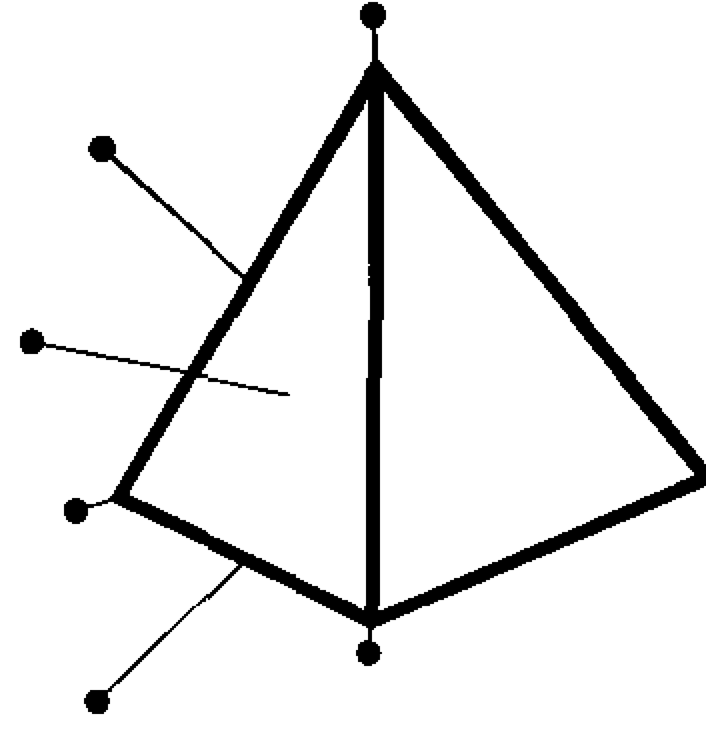
O: the *octahedral group* of order 24 of rotations of a cube, or of a regular octahedron;

I: the *icosahedral group* of 60 rotations of a regular dodecahedron or a regular icosahedron:



We will not attempt to classify the infinite subgroups.

Proof. Let G be a finite subgroup of SO_3, and denote its order by N. Every element g of G except the identity is a rotation about a line ℓ, and this line is obviously unique. So g fixes exactly two points of the unit sphere S in $\mathbb{R}^3$, namely the two points of intersection $\ell \cap S$. We call these points the *poles* of g. Thus a pole is a point p on the unit sphere such that $gp = p$ for some element $g \neq 1$ of G. For example, if G is the group of rotational symmetries of a tetrahedron Δ, then the poles will be the points of S lying over the vertices, the centers of faces, and the centers of edges of Δ.



Let P denote the set of all poles.

(9.2) Lemma The set P is carried to itself by the action of G on the sphere. So G operates on P.

Proof. Let p be a pole, say the pole of $g \in G$. Let x be an arbitrary element of G. We have to show that xp is a pole, meaning that xp is left fixed by some element g' of G other than the identity. The required element is xgx^{-1}: $xgx^{-1}(xp) = xgp = xp$, and $xgx^{-1} \neq 1$ because $g \neq 1$. □

We are now going to get information about the group by counting the poles. Since every element of G except 1 has two poles, our first guess might be that there are $2N - 2$ poles altogether. This isn't quite correct, because the same point p may be a pole for more than one group element.

The stabilizer of a pole p is the group of all of the rotations about the line $\ell = (0, p)$ which are in G. This group is cyclic and is generated by the rotation of smallest angle θ in G. [See the proof of Theorem (3.4a).] If the order of the stabilizer is r_p, then $\theta = 2\pi/r_p$.

We know that $r_p > 1$ because, since p is a pole, the stabilizer G_p contains an element besides 1. By the Counting Formula (7.2),

$$|G_p|\,|O_p| = |G|.$$

We write this equation as

$$r_p n_p = N, \tag{9.3}$$

where n_p is the number of poles in the orbit O_p of p.

The set of elements of G with a given pole p is the stabilizer G_p, minus the identity element. So there are $(r_p - 1)$ group elements with p as pole. On the other hand, every group element g except 1 has two poles. Having to subtract 1 everywhere is a little confusing here, but the correct relation is

$$\sum_{p \in P} (r_p - 1) = 2N - 2. \tag{9.4}$$

Now if p and p' are in the same orbit, then the stabilizers G_p and $G_{p'}$ have the same order. This is because $O_p = O_{p'}$ and $|G| = |G_p|\,|O_p| = |G_{p'}|\,|O_{p'}|$. Therefore we can collect together the terms on the left side of (9.4) which correspond to poles in a given orbit O_p. There are n_p such terms, so the number of poles col-

lected together is $n_p(r_p - 1)$. Let us number the orbits in some way, as $O_1, O_2, \ldots$. Then

$$\sum_i n_i(r_i - 1) = 2N - 2,$$

where $n_i = |O_i|$, and $r_i = |G_p|$ for any $p \in O_i$. Since $N = n_i r_i$, we can divide both sides by N and switch sides, to get the famous formula

$$2 - \frac{2}{N} = \sum_i \left(1 - \frac{1}{r_i}\right). \tag{9.5}$$

This formula may not look very promising at first glance, but actually it tells us a great deal. The left side is less than 2, while each term on the right is at least $\frac{1}{2}$. It follows that there can be at most three orbits!

The rest of the classification is made by listing the various possibilities:

One orbit: $2 - \frac{2}{N} = 1 - \frac{1}{r}$. This is impossible, because $2 - \frac{2}{N} \geq 1$, while $1 - \frac{1}{r} < 1$.

Two orbits: $2 - \frac{2}{N} = \left(1 - \frac{1}{r_1}\right) + \left(1 - \frac{1}{r_2}\right)$, that is, $\frac{2}{N} = \frac{1}{r_1} + \frac{1}{r_2}$.

We know that $r_i \leq N$, because r_i divides N. This equation can hold only if $r_1 = r_2 = N$. Thus $n_1 = n_2 = 1$. There are two poles p, p', both fixed by every element of the group. Obviously, G is the cyclic group C_N of rotations about the line ℓ through p and p'.

Three orbits: This is the main case: Formula (9.5) reduces to

$$\frac{2}{N} = \frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3} - 1.$$

We arrange the r_i in increasing order. Then $r_1 = 2$. For if all r_i were at least 3, then the right side would be ≤ 0, which is impossible.

Case 1: At least two of the orders r_i are 2: $r_1 = r_2 = 2$. The third order $r_3 = r$ can be arbitrary, and $N = 2r$. Then $n_3 = 2$: There is one pair of poles $\{p, p'\}$ making the orbit O_3. Every element g either fixes p and p' or interchanges them. So the elements of G are rotations about $\ell = (p, p')$, or else they are rotations by π about a line ℓ' perpendicular to ℓ. It is easily seen that G is the group of rotations fixing a regular r-gon Δ, the dihedral group D_r. The polygon Δ lies in the plane perpendicular to ℓ, and the vertices and the centers of faces of Δ corresponding to the remaining poles. The bilateral (reflection) symmetries of the polygon in $\mathbb{R}^2$ have become rotations through the angle π when Δ is put into $\mathbb{R}^3$.

Case 2: Only one r_i is 2: The triples $r_1 = 2$, $r_2 \geq 4$, $r_3 \geq 4$ are impossible, because $1/2 + 1/4 + 1/4 - 1 = 0$. Similarly, $r_i = 2$, $r_2 = 3$, $r_3 \geq 6$ can not occur because $1/2 + 1/3 + 1/6 - 1 = 0$. There remain only three possibilities:

(9.6)

(i) $r_i = (2, 3, 3)$, $N = 12$;
(ii) $r_i = (2, 3, 4)$, $N = 24$;
(iii) $r_i = (2, 3, 5)$, $N = 60$.

It remains to analyze these three cases. We will indicate the configurations briefly.

(9.7)

(i) $n_i = (6, 4, 4)$. The poles in the orbit O_2 are the vertices of a regular tetrahedron Δ, and G is the group of rotations fixing it: $G = T$. Here n_1 is the number of edges of Δ, and n_2, n_3 are the numbers of vertices and faces of Δ.
(ii) $n_i = (12, 8, 6)$. The poles in O_2 are the vertices of a cube, and the poles in O_3 are the vertices of a regular octahedron. $G = O$ is the group of their rotations. The integers n_i are the numbers of edges, vertices, and faces of a cube.
(iii) $n_i = (30, 20, 12)$. The poles of O_2 are the vertices of a regular dodecahedron, and those in O_3 are the vertices of a regular icosahedron: $G = I$.

There is still some work to be done to prove the assertions of (9.7). Intuitively, the poles in an orbit should be the vertices of a regular polyhedron because they form a single orbit and are therefore evenly spaced on the sphere. However this is not quite accurate, because the centers of the edges of a cube, for example, form a single orbit but do not span a regular polyhedron. (The figure they span is called a *truncated* polyhedron.)

As an example, consider (9.7iii). Let p be one of the 12 poles in O_3, and let q be one of the poles of O_2 nearest to p. Since the stabilizer of p is of order 5 and operates on O_2 (because G does), the images of q provide a set of five nearest neighbors to p, the poles obtained from q by the five rotations about p in G. Therefore the number of poles of O_2 nearest to p is a multiple of 5, and it is easily seen that 5 is the only possibility. So these five poles are the vertices of a regular pentagon. The 12 pentagons so defined form a regular dodecahedron. □

We close this chapter by remarking that our discussion of the motions of the plane has analogues for the group M_3 of rigid motions of 3-space. In particular, one can define the notion of *crystallographic group*, which is a discrete subgroup whose translation group is a three-dimensional lattice L. To say that L is a lattice means that there are three linearly independent vectors a, b, c in $\mathbb{R}^3$ such that $t_a, t_b, t_c, \in G$. The crystallographic groups are analogous to lattice groups in $M = M_2$, and crystals form examples of three-dimensional configurations having

such groups as symmetry. We imagine the crystal to be infinitely large. Then the fact that the molecules are arranged regularly implies that they form an array having three independent translational symmetries. It has been shown that there are 230 types of crystallographic groups, analogous to the 17 lattice groups (4.15). This is too long a list to be very useful, and so crystals have been classified more crudely into seven *crystal systems*. For more about this, and for a discussion of the 32 crystallographic point groups, look in a book on crystallography.

Un bon héritage vaut mieux que le plus joli problème de géométrie,
parce qu'il tient lieu de méthode générale,
et sert à resoudre bien des problèmes.

Gottfried Wilhelm Leibnitz

EXERCISES

1. Symmetry of Plane Figures

1. Prove that the set of symmetries of a figure F in the plane forms a group.
2. List all symmetries of **(a)** a square and **(b)** a regular pentagon.
3. List all symmetries of the following figures.
   (a) (1.4) **(b)** (1.5) **(c)** (1.6) **(d)** (1.7)
4. Let G be a finite group of rotations of the plane about the origin. Prove that G is cyclic.

2. The Group of Motions of the Plane

1. Compute the fixed point of $t_a\rho_\theta$ algebraically.
2. Verify the rules (2.5) by explicit calculation, using the definitions (2.3).
3. Prove that **O** is not a normal subgroup of M.
4. Let m be an orientation-reversing motion. Prove that m^2 is a translation.
5. Let SM denote the subset of orientation-preserving motions of the plane. Prove that SM is a normal subgroup of M, and determine its index in M.
6. Prove that a linear operator on $\mathbb{R}^2$ is a reflection if and only if its eigenvalues are 1 and -1, and its eigenvectors are orthogonal.
7. Prove that a conjugate of a reflection or a glide reflection is a motion of the same type, and that if m is a glide reflection then the glide vectors of m and of its conjugates have the same length.
8. Complete the proof that (2.13) is a homomorphism.
9. Prove that the map $M \longrightarrow \{1, r\}$ defined by $t_a\rho_\theta \rightsquigarrow 1$, $t_a\rho_\theta r \rightsquigarrow r$ is a homomorphism.
10. Compute the effect of rotation of the axes through an angle η on the expressions $t_a\rho_\theta$ and $t_a\rho_\theta r$ for a motion.

11. (a) Compute the eigenvalues and eigenvectors of the linear operator $m = \rho_\theta r$.
(b) Prove algebraically that m is a reflection about a line through the origin, which subtends an angle of $\frac{1}{2}\theta$ with the x-axis.
(c) Do the same thing as in (b) geometrically.

12. Compute the glide vector of the glide $t_a \rho_\theta r$ in terms of a and θ.

13. (a) Let m be a glide reflection along a line ℓ. Prove geometrically that a point x lies on ℓ if and only if x, $m(x)$, $m^2(x)$ are colinear.
(b) Conversely, prove that if m is an orientation-reversing motion and x is a point such that x, $m(x)$, $m^2(x)$ are distinct points on a line ℓ, then m is a glide reflection along ℓ.

14. Find an isomorphism from the group SM to the subgroup of $GL_2(\mathbb{C})$ of matrices of the form $\begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix}$, with $|a| = 1$.

15. (a) Write the formulas for the motions (2.3) in terms of the complex variable $z = x + iy$.
(b) Show that every motion has the form $m(z) = \alpha z + \beta$ or $m(z) = \alpha \bar{z} + \beta$, where $|\alpha| = 1$ and β is an arbitrary complex number.

3. *Finite Groups of Motions*

1. Let D_n denote the dihedral group (3.6). Express the product $x^2yx^{-1}y^{-1}x^3y^3$ in the form x^iy^j in D_n.

2. List all subgroups of the group D_4, and determine which are normal.

3. Find all proper normal subgroups and identify the quotient groups of the groups D_{13} and D_{15}.

4. (a) Compute the cosets of the subgroup $H = \{1, x^5\}$ in the dihedral group D_{10} explicitly.
(b) Prove that D_{10}/H is isomorphic to D_5.
(c) Is D_{10} isomorphic to $D_5 \times H$?

5. List the subgroups of $G = D_6$ which do not contain $N = \{1, x^3\}$.

6. Prove that every finite subgroup of M is a conjugate subgroup of one of the standard subgroups listed in Corollary (3.5).

4. *Discrete Groups of Motions*

1. Prove that a discrete group G consisting of rotations about the origin is cyclic and is generated by ρ_θ where θ is the smallest angle of rotation in G.

2. Let G be a subgroup of M which contains rotations about two different points. Prove algebraically that G contains a translation.

3. Let (a, b) be a lattice basis of a lattice L in $\mathbb{R}^2$. Prove that every other lattice basis has the form $(a', b') = (a, b)P$, where P is a 2×2 integer matrix whose determinant is ± 1.

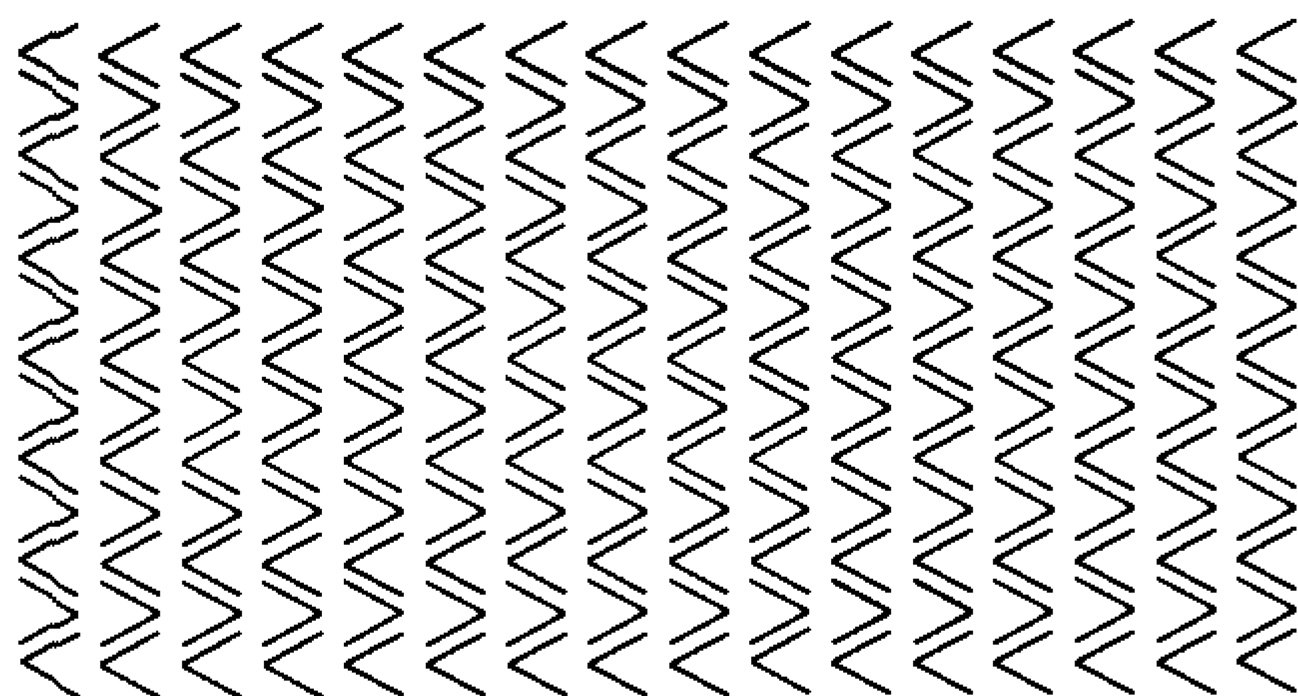
4. Determine the point group for each of the patterns depicted in Figure (4.16).

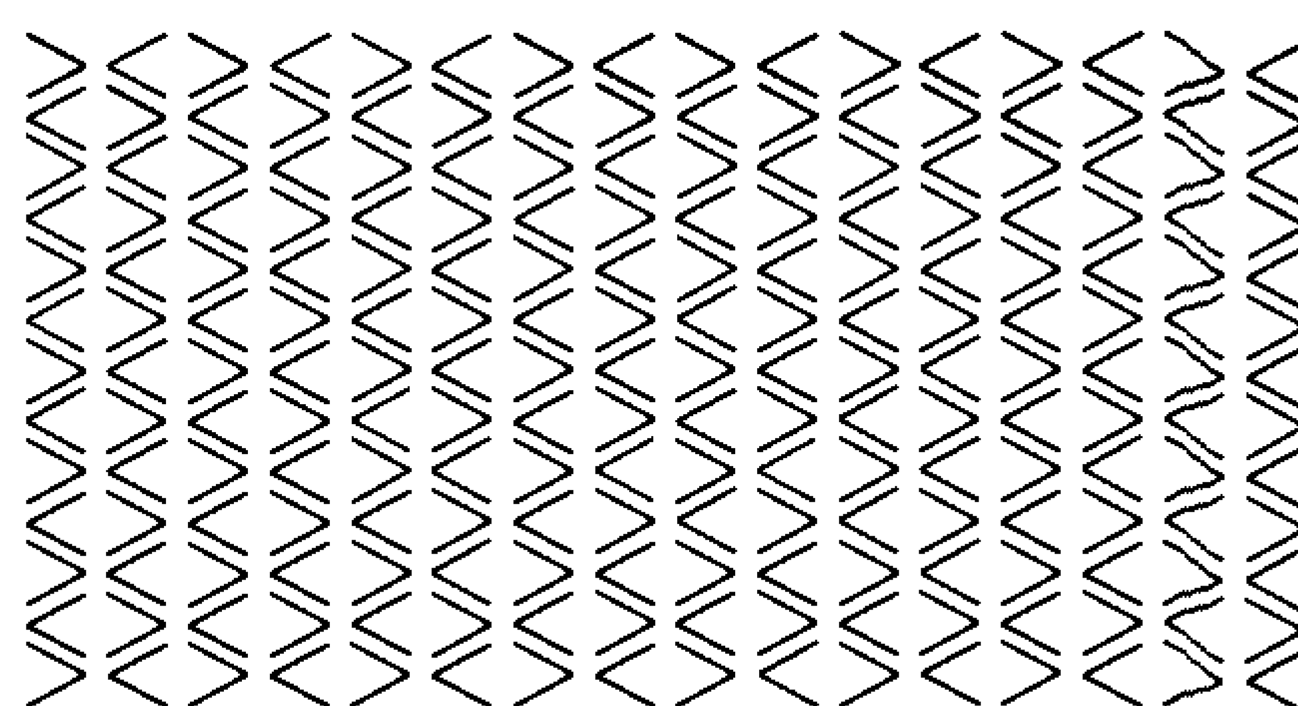
5. (a) Let B be a square of side length a, and let $\epsilon > 0$. Let S be a subset of B such that the distance between any two points of S is $\geq \epsilon$. Find an explicit upper bound for the number of elements in S.
(b) Do the same thing for a box B in $\mathbb{R}^n$.

6. Prove that the subgroup of $\mathbb{R}^+$ generated by 1 and $\sqrt{2}$ is dense in $\mathbb{R}^+$.
7. Prove that every discrete subgroup of $\mathbf{O}$ is finite.
8. Let G be a discrete subgroup of M. Prove that there is a point p_0 in the plane which is not fixed by any point of G except the identity.
9. Prove that the group of symmetries of the frieze pattern

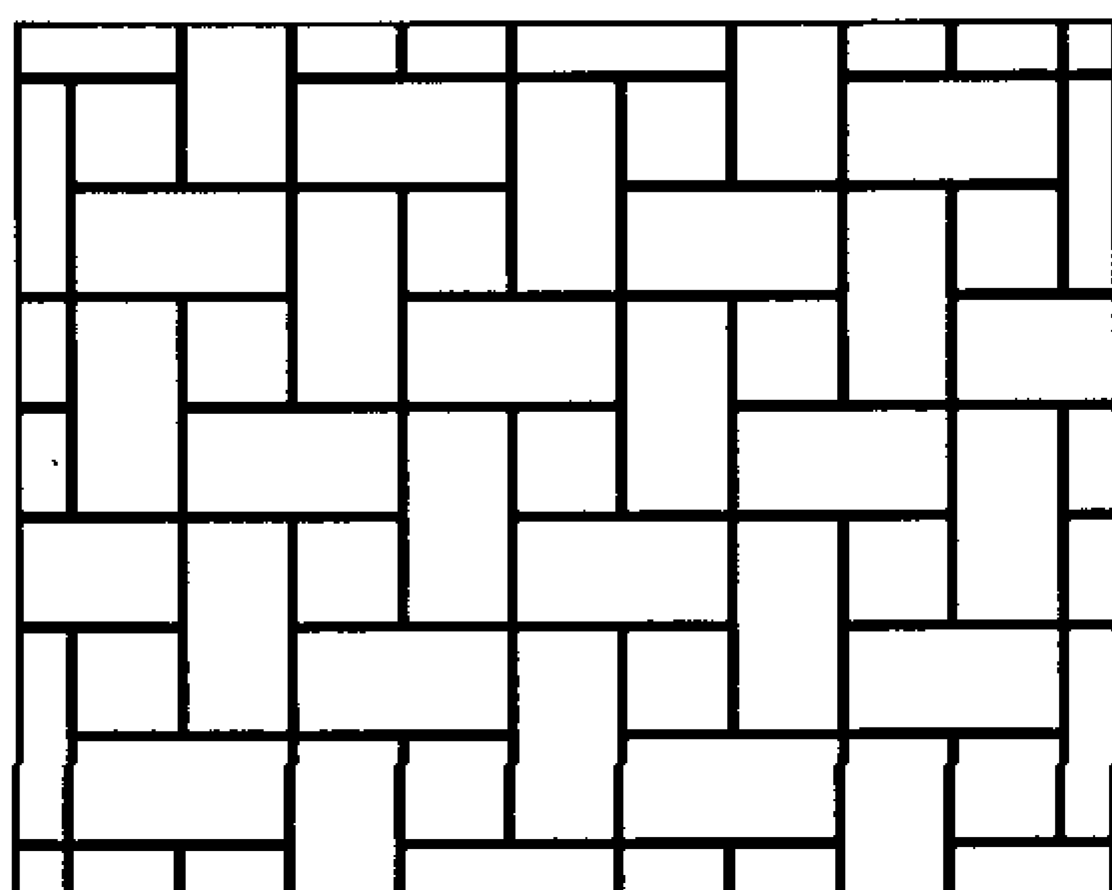
   ...EEEEEEEEEEE...

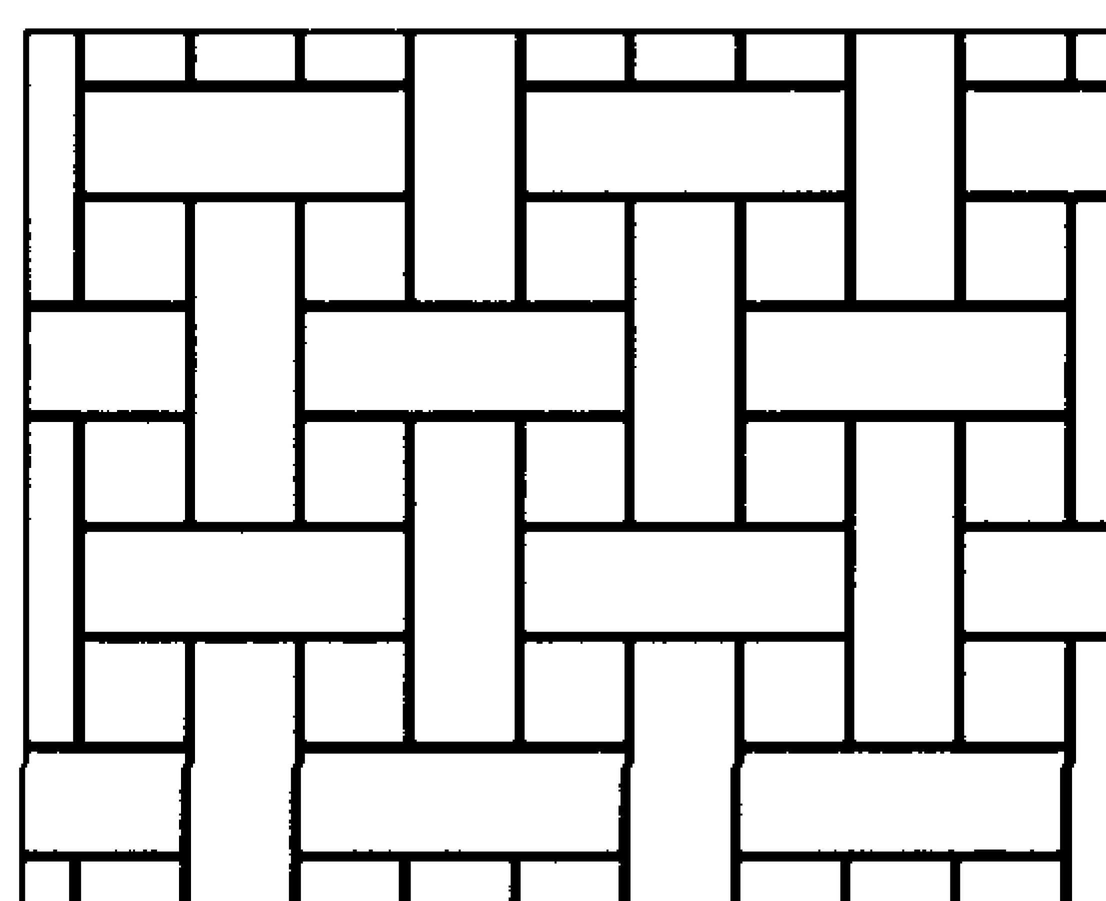
   is isomorphic to the direct product $C_2 \times C_\infty$ of a cyclic group of order 2 and an infinite cyclic group.
10. Let G be the group of symmetries of the frieze pattern ... └┐ ┌┘ └┐ ┌┘ └┐ ┌┘ └┐ ┌┘ ...
    (a) Determine the point group $\overline{G}$ of G.
    (b) For each element $\overline{g} \in \overline{G}$, and each element $g \in G$ which represents $\overline{g}$, describe the action of g geometrically.
    (c) Let H be the subgroup of translations in G. Determine $[G:H]$.
11. Let G be the group of symmetries of the pattern

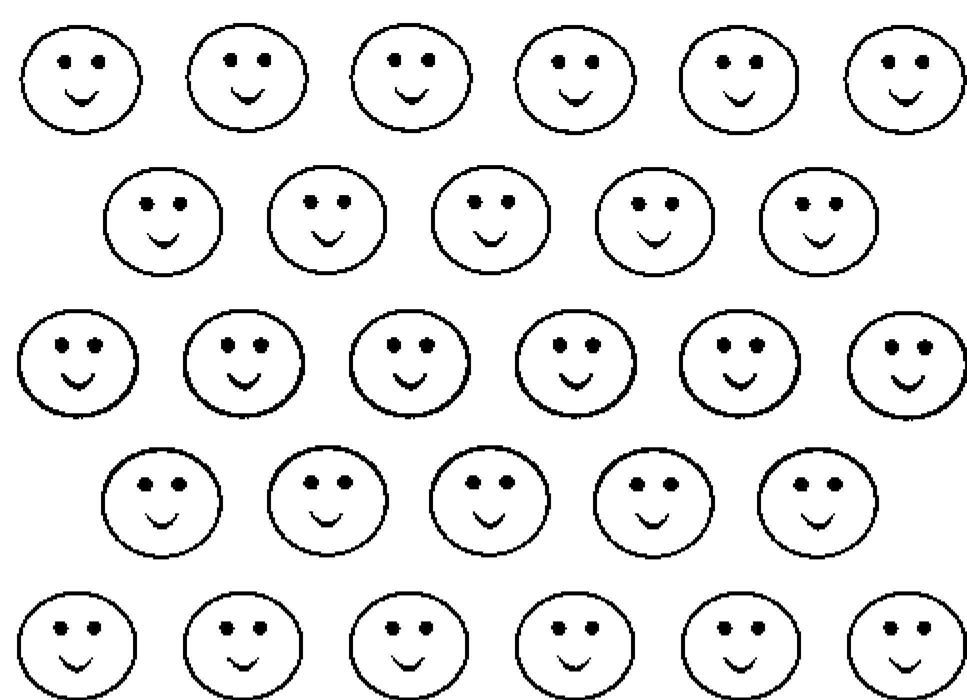


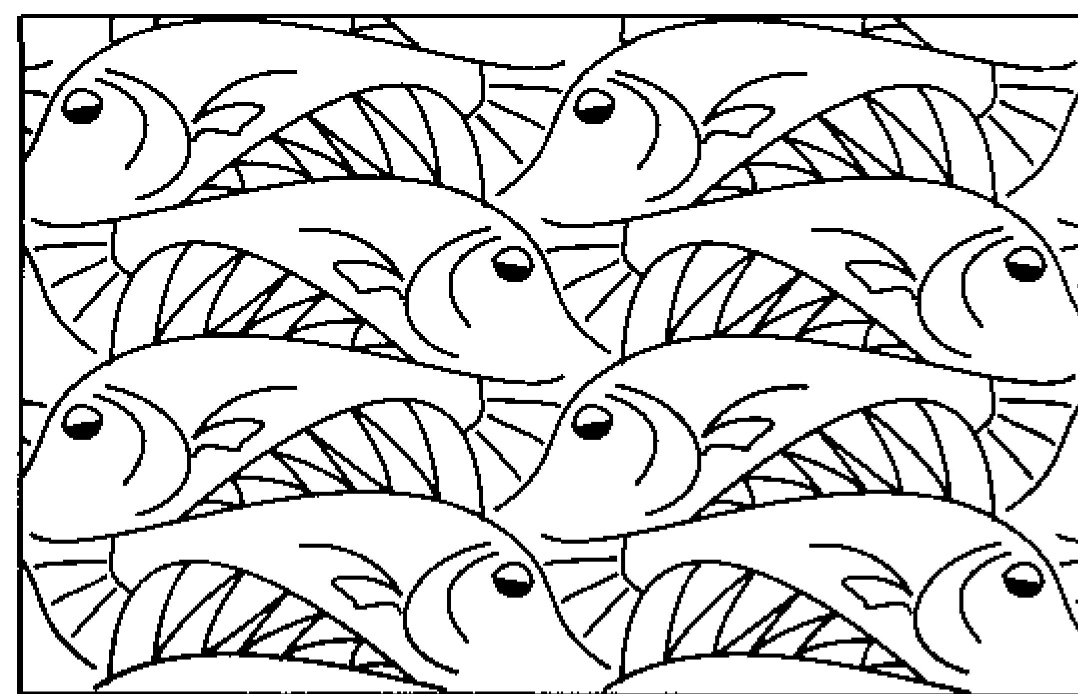


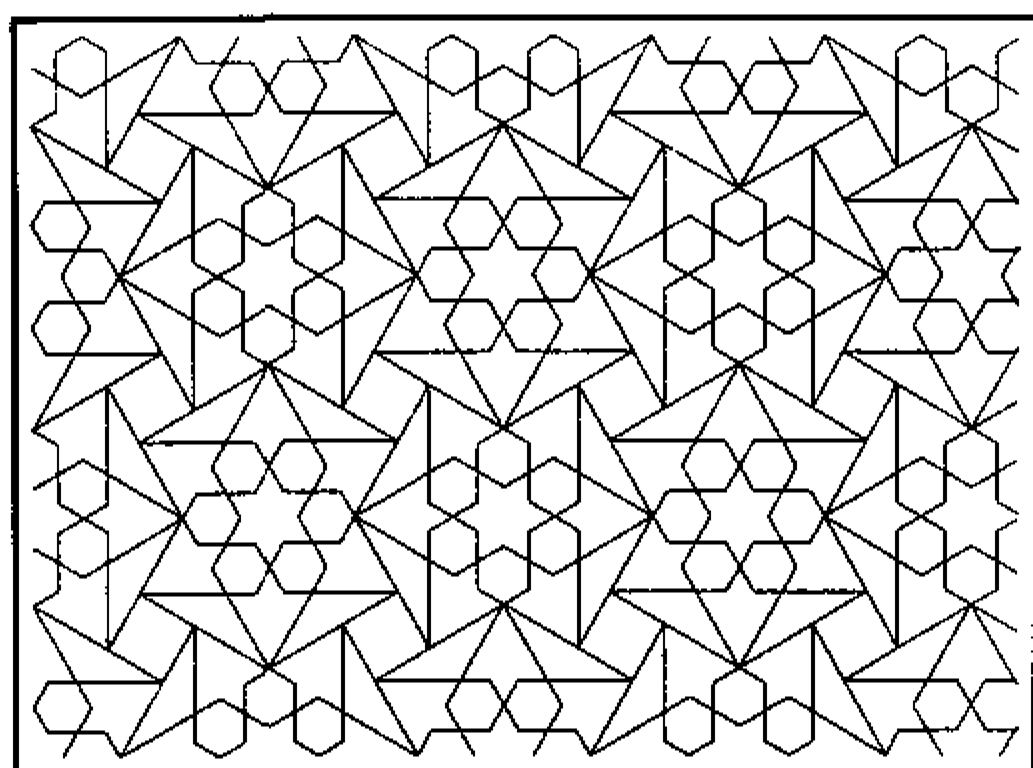
    Determine the point group of G.
12. Let G be the group of symmetries of an equilateral triangular lattice L. Find the index in G of the subgroup $T \cap G$.
13. Let G be a discrete group in which every element is orientation-preserving. Prove that the point group $\overline{G}$ is a cyclic group of rotations and that there is a point p in the plane such that the set of group elements which fix p is isomorphic to $\overline{G}$.
14. With each of the patterns shown, find a pattern with the same type of symmetry in (4.16).

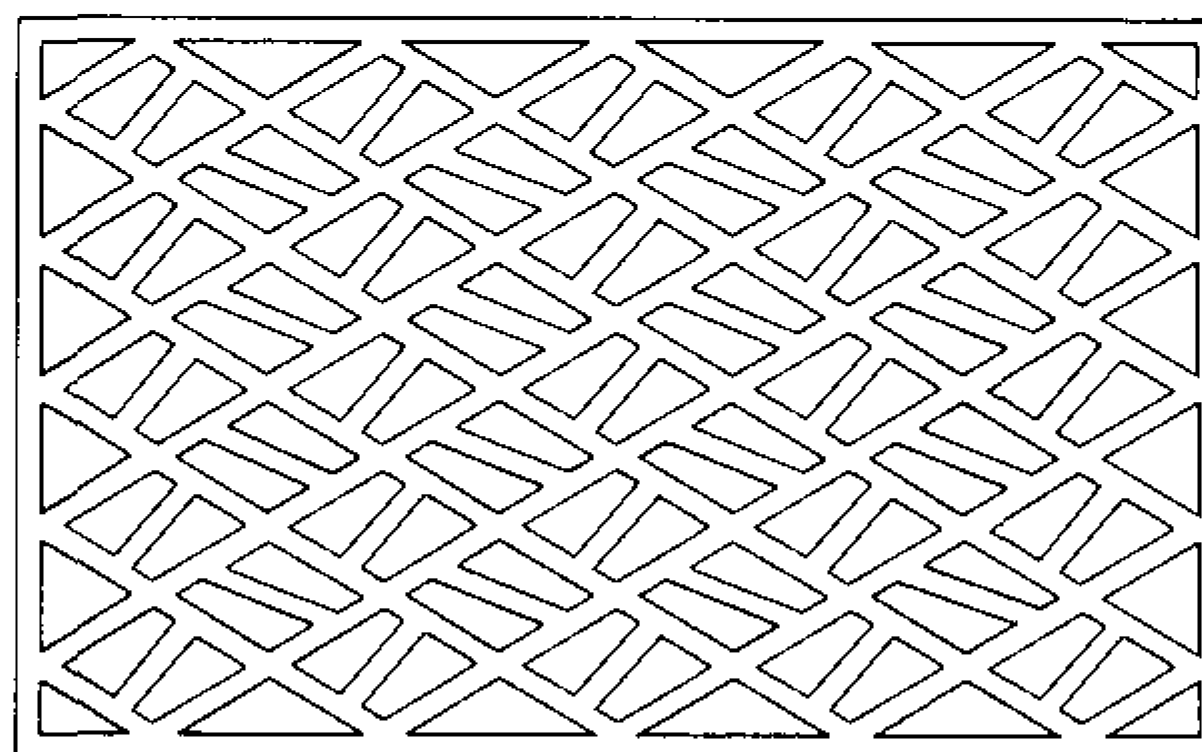












15. Let N denote the group of rigid motions of the line $\ell = \mathbb{R}^1$. Some elements of N are

$$t_a: \longrightarrow x + a, \quad a \in \mathbb{R}, \quad s: x \longrightarrow -x.$$

(a) Show that $\{t_a, t_a s\}$ are all of the elements of N, and describe their actions on ℓ geometrically.

(b) Compute the products $t_a t_b, s t_a, ss$.

(c) Find all discrete subgroups of N which contain a translation. It will be convenient to choose your origin and unit length with reference to the particular subgroup. Prove that your list is complete.

***16.** Let N' be the group of motions of an infinite ribbon

$$R = \{(x,y) \mid -1 \leq y \leq 1\}.$$

It can be viewed as a subgroup of the group M. The following elements are in N':

$$t_a: (x,y) \longrightarrow (x + a, y)$$

$$s: (x,y) \longrightarrow (-x, y)$$

$$r: (x,y) \longrightarrow (x, -y)$$

$$\rho: (x,y) \longrightarrow (-x, -y).$$

(a) Show that these elements generate N', and describe the elements of N' as products.

(b) State and prove analogues of (2.5) for these motions.

(c) A frieze pattern is any pattern on the ribbon which is periodic and not degenerate, in the sense that its group of symmetries is discrete. Since it is periodic, its group of symmetries will contain a translation. Some sample patterns are depicted in the text (1.3, 1.4, 1.6, 1.7). Classify the symmetry groups which arise, identifying those which differ only in the choice of origin and unit length on the ribbon. I suggest that you begin by trying to make patterns with different kinds of symmetry. Please make

a careful case analysis when proving your results. A suitable format would be as follows: Let G be a discrete subgroup containing a translation.

Case 1: Every element of G is a translation. Then ...,

Case 2: G contains the rotation ρ but no orientation-reversing symmetry. Then ..., and so on.

***17.** Let L be a lattice of $\mathbb{R}^2$, and let a, b be linearly independent vectors lying in L. Show that the subgroup $L' = \{ma + nb \mid m, n \in \mathbb{Z}\}$ of L generated by a, b has finite index, and that the index is the number of lattice points in the parallelogram whose vertices are $0, a, b, a + b$ and which are not on the "far edges" $[a, a + b]$ and $[b, a + b]$. (So, 0 is included, and so are points which lie on the edges $[0, a]$, $[0, b]$, except for the points a, b themselves.)

18. (a) Find a subset F of the plane which is not fixed by any motion $m \in M$.
(b) Let G be a discrete group of motions. Prove that the union S of all images of F by elements of G is a subset whose group of symmetries G' contains G.
(c) Show by an example that G' may be larger than G.
*(d) Prove that there exists a subset F such that $G' = G$.

***19.** Let G be a lattice group such that no element $g \neq 1$ fixes any point of the plane. Prove that G is generated by two translations, or else by one translation and one glide.

***20.** Let G be a lattice group whose point group is $D_1 = \{1, r\}$.
(a) Show that the glide lines and the lines of reflection of G are all parallel.
(b) Let $L = L_G$. Show that L contains nonzero vectors $a = (a_1, 0)^t$, $b = (0, b_2)^t$.
(c) Let a and b denote the smallest vectors of the type indicated in (b). Then either (a, b) or (a, c) is a lattice basis for L, where $c = \frac{1}{2}(a + b)$.
(d) Show that if coordinates in the plane are chosen so that the x–axis is a glide line, then G contains one of the elements $g = r$ or $g = t_{\frac{1}{2}a}r$. In either case, show that $G = L \cup Lg$.
(e) There are four possibilities described by the dichotomies (c) and (d). Show that there are only three different kinds of group.

21. Prove that if the point group of a lattice group G is C_6, then $L = L_G$ is an equilateral triangular lattice, and G is the group of all rotational symmetries of L about the origin.

22. Prove that if the point group of a lattice group G is D_6, then $L = L_G$ is an equilateral triangular lattice, and G is the group of all symmetries of L.

***23.** Prove that symmetry groups of the figures in Figure (4.16) exhaust the possibilities.

5. Abstract Symmetry: Group Operations

1. Determine the group of automorphisms of the following groups.
(a) C_4 (b) C_6 (c) $C_2 \times C_2$

2. Prove that (5.4) is an equivalence relation.

3. Let S be a set on which G operates. Prove that the relation $s \sim s'$ if $s' = gs$ for some $g \in G$ is an equivalence relation.

4. Let $\varphi: G \longrightarrow G'$ be a homomorphism, and let S be a set on which G' operates. Show how to define an operation of G on S, using the homomorphism φ.

5. Let $G = D_4$ be the dihedral group of symmetries of the square.
   (a) What is the stabilizer of a vertex? an edge?
   (b) G acts on the set of two elements consisting of the diagonal lines. What is the stabilizer of a diagonal?
6. In each of the figures in exercise 14 of Section 4, find the points which have nontrivial stabilizers, and identify the stabilizers.

*7. Let G be a discrete subgroup of M.
   (a) Prove that the stabilizer G_p of a point p is finite.
   (b) Prove that the orbit O_p of a point p is a discrete set, that is, that there is a number $\epsilon > 0$ so that the distance between two distinct points of the orbit is at least ϵ.
   (c) Let B, B' be two bounded regions in the plane. Prove that there are only finitely many elements $g \in G$ so that $gB \cap B'$ is nonempty.
8. Let $G = GL_n(\mathbb{R})$ operate on the set $S = \mathbb{R}^n$ by left multiplication.
   (a) Describe the decomposition of S into orbits for this operation.
   (b) What is the stabilizer of e_1?
9. Decompose the set $\mathbb{C}^{2\times 2}$ of 2×2 complex matrices for the following operations of $GL_2(\mathbb{C})$:
   (a) Left multiplication
   *(b) Conjugation
10. (a) Let $S = \mathbb{R}^{m\times n}$ be the set of real $m \times n$ matrices, and let $G = GL_m(\mathbb{R}) \times GL_n(\mathbb{R})$. Prove that the rule $(P, Q), A \rightsquigarrow PAQ^{-1}$ defines an operation of G on S.
    (b) Describe the decomposition of S into G-orbits.
    (c) Assume that $m \le n$. What is the stabilizer of the matrix $[I \mid 0]$?
11. (a) Describe the orbit and the stabilizer of the matrix $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ under conjugation in $GL_n(\mathbb{R})$.
    (b) Interpreting the matrix in $GL_2(\mathbb{F}_3)$, find the order (the number of elements) of the orbit.
12. (a) Define automorphism of a field.
    (b) Prove that the field $\mathbb{Q}$ of rational numbers has no automorphism except the identity.
    (c) Determine Aut F, when $F = \mathbb{Q}[\sqrt{2}]$.

6. The Operation on Cosets

1. What is the stabilizer of the coset aH for the operation of G on G/H?
2. Let G be a group, and let H be the cyclic subgroup generated by an element x of G. Show that if left multiplication by x fixes every coset of H in G, then H is a normal subgroup.
3. (a) Exhibit the bijective map (6.4) explicitly, when G is the dihedral group D_4 and S is the set of vertices of a square.
   (b) Do the same for D_n and the vertices of a regular n–gon.
4. (a) Describe the stabilizer H of the index **1** for the action of the symmetric group $G = S_n$ on $\{\mathbf{1}, \ldots, \mathbf{n}\}$ explicitly.
   (b) Describe the cosets of H in G explicitly for this action.
   (c) Describe the map (6.4) explicitly.

5. Describe all ways in which S_3 can operate on a set of four elements.
6. Prove Proposition (6.5).
7. A map $S \longrightarrow S'$ of G-sets is called a *homomorphism* of G-sets if $\varphi(gs) = g\varphi(s)$ for all $s \in S$ and $g \in G$. Let φ be such a homomorphism. Prove the following:
   (a) The stabilizer $G_{\varphi(s)}$ contains the stabilizer G_s.
   (b) The orbit of an element $s \in S$ maps onto the orbit of $\varphi(s)$.

7. The Counting Formula

1. Use the counting formula to determine the orders of the group of rotational symmetries of a cube and of the group of rotational symmetries of a tetrahedron.
2. Let G be the group of rotational symmetries of a cube C. Two regular tetrahedra Δ, Δ' can be inscribed in C, each using half of the vertices. What is the order of the stabilizer of Δ?
3. Compute the order of the group of symmetries of a dodecahedron, when orientation-reversing symmetries such as reflections in planes, as well as rotations, are allowed. Do the same for the symmetries of a cube and of a tetrahedron.
4. Let G be the group of rotational symmetries of a cube, let S_e, S_v, S_f be the sets of vertices, edges, and faces of the cube, and let H_v, H_e, H_f be the stabilizers of a vertex, an edge, and a face. Determine the formulas which represent the decomposition of each of the three sets into orbits for each of the subgroups.
5. Let $G \supset H \supset K$ be groups. Prove the formula $[G : K] = [G : H][H : K]$ without the assumption that G is finite.
6. (a) Prove that if H and K are subgroups of finite index of a group G, then the intersection $H \cap K$ is also of finite index.
   (b) Show by example that the index $[H : H \cap K]$ need not divide $[G : K]$.

8. Permutation Representations

1. Determine all ways in which the tetrahedral group T (see (9.1)) can operate on a set of two elements.
2. Let S be a set on which a group G operates, and let $H = \{g \in G \mid gs = s \text{ for all } s \in S\}$. Prove that H is a normal subgroup of G.
3. Let G be the dihedral group of symmetries of a square. Is the action of G on the vertices a faithful action? on the diagonals?
4. Suppose that there are two orbits for the operation of a group G on a set S, and that they have orders m, n respectively. Use the operation to define a homomorphism from G to the product $S_m \times S_n$ of symmetric groups.
5. A group G operates faithfully on a set S of five elements, and there are two orbits, one of order 3 and one of order 2. What are the possibilities for G?
6. Complete the proof of Proposition (8.2).
7. Let $F = \mathbb{F}_3$. There are four one-dimensional subspaces of the space of column vectors F^2. Describe them. Left multiplication by an invertible matrix permutes these subspaces. Prove that this operation defines a homomorphism $\varphi: GL_2(F) \longrightarrow S_4$. Determine the kernel and image of this homomorphism.

*8. For each of the following groups, find the smallest integer n such that the group has a faithful operation on a set with n elements.
(a) the quaternion group H (b) D_4 (c) D_6

9. Finite Subgroups of the Rotation Group

1. Describe the orbits of poles for the group of rotations of an octahedron and of an icosahedron.
2. Identify the group of symmetries of a baseball, taking the stitching into account and allowing orientation-reversing symmetries.
3. Let O be the group of rotations of a cube. Determine the stabilizer of a diagonal line connecting opposite vertices.
4. Let $G = O$ be the group of rotations of a cube, and let H be the subgroup carrying one of the two inscribed tetrahedra to itself (see exercise 2, Section 7). Prove that $H = T$.
5. Prove that the icosahedral group has a subgroup of order 10.
6. Determine all subgroups of the following groups:
(a) T (b) I
7. Explain why the groups of symmetries of the cube and octahedron, and of the dodecahedron and icosahedron, are equal.

*8. (a) The 12 points $(\pm1, \pm\alpha, 0), (0, \pm1, \pm\alpha)(\pm\alpha, 0, \pm1)$ form the vertices of a regular icosahedron if α is suitably chosen. Verify this, and determine α.
(b) Determine the matrix of the rotation through the angle $2\pi/5$ about the origin in $\mathbb{R}^2$.
(c) Determine the matrix of the rotation of $\mathbb{R}^3$ through the angle $2\pi/5$ about the axis containing the point $(1, \alpha, 0)$.

*9. Prove the crystallographic restriction for three-dimensional crystallographic groups: A rotational symmetry of a crystal has order 2, 3, 4, or 6.

Miscellaneous Problems

1. Describe completely the following groups:
(a) Aut D_4 (b) Aut H, where H is the quaternion group
2. (a) Prove that the set Aut G of automorphisms of a group G forms a group.
(b) Prove that the map φ: $G \longrightarrow$ Aut G defined by $g \rightsquigarrow$ (conjugation by g) is a homomorphism, and determine its kernel.
(c) The automorphisms which are conjugation by a group element are called *inner automorphisms*. Prove that the set of inner automorphisms, the image of φ, is a normal subgroup of Aut G.
3. Determine the quotient group Aut H/Int H for the quaternion group H.

*4. Let G be a lattice group. A *fundamental domain* D for G is a bounded region in the plane, bounded by piecewise smooth curves, such that the sets gD, $g \in G$ cover the plane without overlapping except along the edges. We assume that D has finitely many connected components.
(a) Find fundamental domains for the symmetry groups of the patterns illustrated in exercise 14 of Section 4.
(b) Show that any two fundamental domains D, D' for G can be cut into finitely many congruent pieces of the form $gD \cap D'$ or $D \cap gD'$ (see exercise 7, Section 5).

(c) Conclude that D and D' have the same area. (It may happen that the boundary curves intersect infinitely often, and this raises some questions about the definition of area. Disregard such points in your answer.)

*5. Let G be a lattice group, and let p_0 be a point in the plane which is not fixed by any element of G. Let $S = \{gp_0 \mid g \in G\}$ be the orbit of p_0. The plane can be divided into polygons, each one containing a single point of S, as follows: The polygon Δ_p containing p is the set of points q whose distance from p is the smallest distance to any point of S:

$$\Delta_p = \{q \in \mathbb{R}^2 \mid \text{dist}(q, p) \le \text{dist}(q, p') \text{ for all } p' \in S\}.$$

(a) Prove that Δ_p is a polygon.
(b) Prove that Δ_p is a fundamental domain for G.
(c) Show that this method works for all discrete subgroups of M, except that the domain Δ_p which is constructed need not be a bounded set.
(d) Prove that Δ_p is bounded if and only if the group is a lattice group.

*6. (a) Let $G' \subset G$ be two lattice groups. Let D be a fundamental domain for G. Show that a fundamental domain D' for G' can be constructed out of finitely many translates gD of D.
(b) Show that $[G : G'] < \infty$ and that $[G : G'] = \text{area}(D')/\text{area}(D)$.
(c) Compute the index $[G : L_G]$ for each of the patterns (4.16).

*7. Let G be a finite group operating on a finite set S. For each element $g \in G$, let S^g denote the subset of elements of S fixed by g: $S^g = \{s \in S \mid gs = s\}$.
(a) We may imagine a true–false table for the assertion that $gs = s$, say with rows indexed by elements of G and columns indexed by elements. Construct such a table for the action of the dihedral group D_3 on the vertices of a triangle.
(b) Prove the formula $\sum_{s \in S} |G_s| = \sum_{g \in G} |S^g|$.
(c) Prove *Burnside's Formula*:

$$|G| \cdot (\textit{number of orbits}) = \sum_{g \in G} |S^g|.$$

8. There are $70 = \binom{8}{4}$ ways to color the edges of an octagon, making four black and four white. The group D_8 operates on this set of 70, and the orbits represent equivalent colorings. Use Burnside's Formula to count the number of equivalence classes.

9. Let G be a group of order n which operates nontrivially on a set of order r. Prove that if $n > r!$, then G has a proper normal subgroup.

Chapter 6

More Group Theory

The more to do or to prove, the easier the doing or the proof.

James Joseph Sylvester

1. THE OPERATIONS OF A GROUP ON ITSELF

By an operation of a group G on itself, we mean that in the definition of the operation, G plays the role both of the group and of the set on which it operates. Any group operates on itself in several ways, two of which we single out here. The first is *left multiplication:*

(1.1)
$$G \times G \longrightarrow G$$
$$g, x \rightsquigarrow gx.$$

This is obviously a transitive operation of G on G, that is, G forms a single orbit, and the stabilizer of any element is the identity subgroup $\{1\}$. So the action is faithful, and the homomorphism

(1.2)
$$G \longrightarrow \text{Perm}(G)$$
$$g \rightsquigarrow m_g = \text{left multiplication by } g$$

defined in Chapter 5, Section 8 is injective.

(1.3) **Theorem.** *Cayley's Theorem:* Every finite group G is isomorphic to a subgroup of a permutation group. If G has order n, then it is isomorphic to a subgroup of the symmetric group S_n.

Proof. Since the operation by left multiplication is faithful, G is isomorphic to its image in $\text{Perm}(G)$. If G has order n, then $\text{Perm}(G)$ is isomorphic to S_n. □

Though Cayley's Theorem is intrinsically interesting, it is not especially useful for computation because S_n, having order $n!$, is too large in comparison with n.

The second operation we will consider is more subtle. It is *conjugation*, the map $G \times G \longrightarrow G$, defined by

$$(g, x) \rightsquigarrow gxg^{-1}. \tag{1.4}$$

For obvious reasons, we will not use multiplicative notation for this operation. You should verify the axioms (5.1) in Chapter 5, introducing a temporary notation such as $g * x$ to denote the conjugate gxg^{-1}.

The stabilizer of an element $x \in G$ for the operation of conjugation has a special name. It is called the *centralizer* of x and is denoted by $Z(x)$:

$$Z(x) = \{g \in G \mid gxg^{-1} = x\} = \{g \in G \mid gx = xg\}. \tag{1.5}$$

The centralizer is the set of group elements which commute with x. Note that $x \in Z(x)$, because x commutes with itself.

The orbit of x for the operation of conjugation is called the *conjugacy class* of x. It consists of all conjugate elements gxg^{-1}. We often write the conjugacy class as

$$C_x = \{x' \in G \mid x' = gxg^{-1} \text{ for some } g \in G\}. \tag{1.6}$$

By the Counting Formula [Chapter 5 (7.2)], $|G| = |C_x|\,|Z(x)|$.

Since the conjugacy classes are orbits for a group operation, they partition G. This gives us what is called the *Class Equation* for a finite group [see Chapter 5(7.3)]:

$$|G| = \sum_{\substack{\text{conjugacy}\\ \text{classes } C}} |C|. \tag{1.7}$$

If we number the conjugacy classes, say as C_i, $i = 1, \ldots, k$, then this formula reads

$$|G| = |C_1| + \cdots + |C_k|.$$

However there is some danger of confusion, because the subscript i in C_i is an index, while the notation C_x as used above stands for the conjugacy class containing the element x of G. In particular, C_1 has two meanings. Perhaps it will be best to list the conjugacy class of the identity element 1 of G first. Then the two interpretations of C_1 will agree.

Notice that the identity element is left fixed by all $g \in G$. Thus C_1 consists of the element 1 alone. Note also that each term on the right side of (1.7), being the order of an orbit, divides the left side. This is a strong restriction on the combinations of integers which may occur in such an equation.

(1.8) *The numbers on the right side of the Class Equation divide the order of the group, and at least one of them is equal to* 1.

For example, the conjugacy classes in the dihedral group D_3, presented as in Chapter 5 (3.6), are the following three subsets:

$$\{1\}, \{x, x^2\}, \{y, xy, x^2y\}.$$

The two rotations x, x^2 are conjugate, as are the three reflections. The Class Equation for D_3 is

$$6 = 1 + 2 + 3. \tag{1.9}$$

Recall from Chapter 2 (4.10) that the center of a group G is the set Z of elements which commute with all elements of the group:

$$Z = \{g \in G \mid gx = xg \text{ for all } x \in G\}.$$

Now the conjugacy class of an element x consists of that element alone if and only if $x = gxg^{-1}$ for all $g \in G$. This means that x is in the center. Thus the elements of the center are represented by 1 on the right side of the Class Equation.

The next proposition follows directly from the definitions.

(1.10) **Proposition.** An element x is in the center of a group G if and only if its centralizer $Z(x)$ is the whole group. □

One case in which the Class Equation (1.7) can be used effectively is when the order of G is a positive power of a prime p. Such a group is called a *p-group*. Here are a few applications of the Class Equation to p-groups.

(1.11) **Proposition.** The center of a p-group G has order > 1.

Proof. The left side of (1.7) is a power of p, say p^e. Also, every term on the right side is a power of p too, because it divides p^e. We want to show that some group element $x \neq 1$ is in the center, which is the same as saying that more than one term on the right side of (1.7) is equal to 1. Now the terms other than 1, being positive powers of p, are divisible by p. Suppose that the class C_1 made the only contribution of 1 to the right side. Then the equation would read

$$p^e = 1 + \sum(\textit{multiples of } p),$$

which is impossible unless $e = 0$. □

The argument used in this proof can be turned around and abstracted to give the following important *Fixed Point Theorem* for actions of p-groups:

(1.12) **Proposition.** Let G be a p-group, and let S be a finite set on which G operates. Assume that the order of S is not divisible by p. Then there is a fixed point for the action of G on S, that is, an element $s \in S$ whose stabilizer is the whole group. □

(1.13) **Proposition.** Every group of order p^2 is abelian.

Proof. Let G be a group of order p^2. We will show that for every $x \in G$, the centralizer $Z(x)$ is the whole group. Proposition (1.10) will then finish the proof. So let $x \in G$. If x is in the center Z, then $Z(x) = G$ as claimed. If $x \notin Z$, then $Z(x)$ is strictly larger than Z, because it contains Z and also contains the element x. Now the orders of Z and $Z(x)$ divide $|G| = p^2$, and Proposition (1.11) tells us that $|Z|$ is at

least p. The only possibility is that $|Z(x)| = p^2$. Hence $Z(x) = G$, and x was in the center after all. □

There are nonabelian groups of order p^3. The dihedral group D_4, for example, has order 8.

Let us use (1.13) to classify groups of order p^2.

(1.14) **Corollary.** Every group of order p^2 is of one of the following types:

(i) a cyclic group of order p^2;

(ii) a product of two cyclic groups of order p.

Proof. Since the order of an element divides p^2, there are two cases to consider:

Case 1: G contains an element of order p^2 and is therefore a cyclic group.

Case 2: Every element x of G except the identity has order p. Let x, y be two elements different from 1, and let H_1, H_2 be the cyclic groups of order p generated by x and y respectively. We may choose y so that it is not a power of x. Then since $y \notin H_1$, $H_1 \cap H_2$ is smaller than H_2, which has order p. So $H_1 \cap H_2 = \{1\}$. Also, the subgroups H_i are normal because G is abelian. Since $y \notin H_1$, the group H_1H_2 is strictly larger than H_1, and its order divides p^2. Thus $H_1H_2 = G$. By Chapter 2 (8.6), $G \approx H_1 \times H_2$. □

The number of possibilities for groups of order p^n increases rapidly with n. There are five isomorphism classes of groups of order 8, and 14 classes of groups of order 16.

2. THE CLASS EQUATION OF THE ICOSAHEDRAL GROUP

In this section we determine the conjugacy classes in the icosahedral group I of rotational symmetries of a dodecahedron, and use them to study this very interesting group. As we have seen, the order of the icosahedral group is 60. It contains rotations by multiples of $2\pi/5$ about the centers of the faces of the dodecahedron, by multiples of $2\pi/3$ about the vertices, and by π about the centers of the edges. Each of the 20 vertices has a stabilizer of order 3, and opposite vertices have the same stabilizer. Thus there are 10 subgroups of order 3—the stabilizers of the vertices. Each subgroup of order 3 contains two elements of order 3, and the intersection of any two of these subgroups consists of the identity element alone. So I contains $10 \times 2 = 20$ elements of order 3. Similarly, the faces have stabilizers of order 5, and there are six such stabilizers, giving us $6 \times 4 = 24$ elements of order 5. There are 15 stabilizers of edges, and these stabilizers have order 2. So there are 15 elements of order 2. Finally, there is one element of order 1. Since

$$60 = 1 + 15 + 20 + 24, \tag{2.1}$$

we have listed all elements of the group.

Equation (2.1) is obtained by partitioning the group according to the orders of the elements. It is closely related to the Class Equation, but we can see that (2.1) is not the Class Equation itself, because 24, which appears on the right side, does not divide 60. On the other hand, we do know that conjugate elements have the same order. So the Class Equation is obtained by subdividing this partition of G still further. Also, note that the *subgroups* of order 3 are all conjugate. This is a general property of group operations, because they are the stabilizers of the vertices, which form a single orbit [Chapter 5 (6.5)]. The same is true for the subgroups of order 5 and for those of order 2.

Clearly the 15 elements of order 2, being the nontrivial elements in conjugate subgroups of order 2, form one conjugacy class. What about the elements of order 3? Let x denote a counterclockwise rotation by $2\pi/3$ about a vertex v. Though x will be conjugate to rotation with the same angle about any other vertex [Chapter 5 (6.5)], it is not so clear whether or not x is conjugate to x^2. Perhaps the first guess would be that x and x^2 are not conjugate.

Let v' denote the vertex opposite to v, and let x' be the counterclockwise rotation by $2\pi/3$ about v'. So x and x' are conjugate elements of the group. Notice that the counterclockwise rotation x about v is the same motion as the clockwise rotation by $2\pi/3$ about the opposite vertex v'. Thus $x^2 = x'$, and this shows that x and x^2 are conjugate after all. It follows that all the elements of order 3 are conjugate. Similarly, the 12 rotations by $2\pi/5$ and $-2\pi/5$ are conjugate. They are not conjugate to the remaining 12 rotations by $4\pi/5$, $-4\pi/5$ of order 5. (One reason, as we have already remarked, is that the order of a conjugacy class divides the order of the group, and 24 does not divide 60.) Thus there are two conjugacy classes of elements of order 5, and the Class Equation is

$$60 = 1 + 15 + 20 + 12 + 12. \tag{2.2}$$

We will now use this Class Equation to prove the following theorem.

(2.3) **Theorem.** The icosahedral group I has no proper normal subgroup.

A group $G \neq \{1\}$ is called a *simple group* if it is not the trivial group and if it contains no proper normal subgroup (no normal subgroup other than $\{1\}$ and G). Thus the theorem can be restated as follows:

(2.4) *The icosahedral group is a simple group.*

Cyclic groups of prime order contain no proper subgroup at all and are therefore simple groups. All other groups, except for the trivial group, contain proper subgroups, though not necessarily normal ones. We should emphasize that this use of the word *simple* does not imply "uncomplicated." Its meaning here is roughly "not compound."

Proof of Theorem (2.3). The proof of the following lemma is straightforward:

(2.5) **Lemma.**

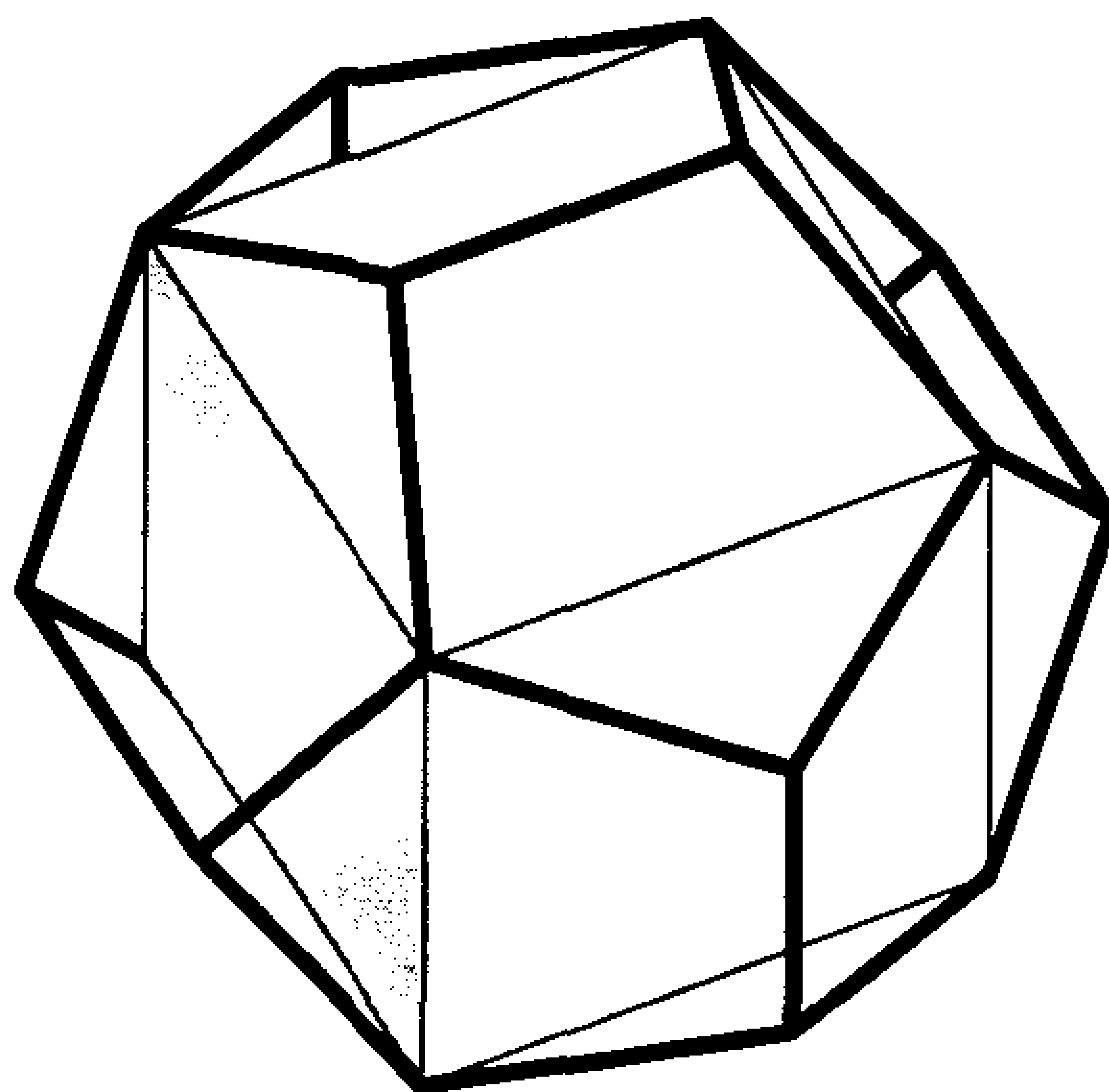
(a) If a normal subgroup N of a group G contains an element x, then it contains the conjugacy class C_x of x in G. In other words, a normal subgroup is a union of conjugacy classes.

(b) The order of a normal subgroup N of G is the sum of the orders of the conjugacy classes which it contains. □

We now apply this lemma. The order of a proper normal subgroup of the icosahedral group is a proper divisor of 60 and is also the sum of some of the terms on the right side of the Class Equation (2.2), including the term 1. It happens that there is no such integer. This proves the theorem. □

(2.6) **Theorem.** The icosahedral group is isomorphic to the alternating group A_5.

Proof. To describe this isomorphism, we need to find a set S of five elements on which I operates. One such set consists of the five cubes which can be inscribed into a dodecahedron, one of which is illustrated below:



(2.7) **Figure.** One of the cubes inscribed in a dodecahedron.

The group I operates on this set of cubes S, and this operation defines a homomorphism $\varphi: I \longrightarrow S_5$, the associated permutation representation. The map φ is our isomorphism from I to its image A_5. To show that it is an isomorphism, we will use the fact that I is a simple group, but we need very little information about the operation itself.

Since the kernel of φ is a normal subgroup of I and since I is a simple group, $\ker\varphi$ is either $\{1\}$ or I. To say $\ker\varphi = I$ would mean that the operation of I on the set of five cubes was the trivial operation, which it is not. Therefore $\ker\varphi = \{1\}$, and φ is injective, defining an isomorphism of I onto its image in S_5.

Let us denote the image in S_5 by I too. We restrict the sign homomorphism $S_5 \longrightarrow \{\pm 1\}$ to I, obtaining a homomorphism $I \longrightarrow \{\pm 1\}$. If this homomorphism were surjective, its kernel would be a normal subgroup of I of order 30 [Chapter 2 (6.15)]. This is impossible because I is simple. Therefore the restriction is the trivial

homomorphism, which just means that I is contained in the kernel A_5 of the sign homomorphism. Since both groups have order 60, $I = A_5$. □

3. OPERATIONS ON SUBSETS

Whenever a group G operates on a set S, there is also an operation on subsets. If $U \subset S$ is a subset, then

(3.1) $$gU = \{gu \mid u \in U\}$$

is another subset of S. The axioms for an operation are clearly verified. So G operates on the set of subsets of S. We can consider the operation on subsets of a given order if we want to do so. Since multiplication by g is a permutation of S, the subsets U and gU have the same order.

For example, let O be the octahedral group of 24 rotations of a cube, and let S be the set of vertices of the cube. Consider the operation of O on subsets of order 2 of S, that is, on unordered pairs of vertices. There are 28 such pairs, and they form three orbits for the group:

(i) {pairs of vertices on an edge};
(ii) {pairs which are opposite on a face of the cube};
(iii) {pairs which are opposite on the cube}.

These orbits have orders 12, 12, and 4 respectively: $28 = 12 + 12 + 4$.

The stabilizer of a subset U is the set of group elements g such that $gU = U$. Thus the stabilizer of a pair of opposite vertices on a face contains two elements—the identity and the rotation by π about the face. This agrees with the counting formula: $24 = 2 \cdot 12$.

Note this important point once more: The equality $gU = U$ does not mean that g leaves the elements in U fixed, but rather that g permutes the elements within U, that is, that $gu \in U$ whenever $u \in U$.

(3.2) **Proposition.** Let H be a group which operates on a set S, and let U be a subset of S. Then H stabilizes U if and only if U is a union of H-orbits. □

This proposition just restates the fact that the H-orbit of an element $u \in U$ is the set of all elements hu. If H stabilizes U, then U contains the H-orbit of any of its elements. □

Let's consider the case that G operates by left multiplication on the subsets of G. Any subgroup H of G is a subset, and its orbit consists of the left cosets. This operation of G on cosets was defined in Chapter 5 (6.1). But any subset of G has an orbit.

(3.3) **Example.** Let $G = D_3$ be the dihedral group of symmetries of an equilateral triangle, presented as usual:

$$G = \{x^i y^j \mid 0 \le i \le 2, 0 \le j \le 1, x^3 = 1, y^2 = 1, yx = x^2 y\}.$$

This group contains 15 subsets of order 2, and we can decompose this set of 15 into orbits for left multiplication. There are three subgroups of order 2:

(3.4) $$H_1 = \{1, y\}, \quad H_2 = \{1, xy\}, \quad H_3 = \{1, x^2y\}.$$

Their cosets form three orbits of order 3. The other six subsets of order 2 form a single orbit: $15 = 3 + 3 + 3 + 6$. The orbit of six is

(3.5) $$\{1, x\}, \{x, x^2\}, \{x^2, 1\}, \{y, x^2y\}, \{xy, y\}, \{x^2y, xy\}. \square$$

(3.6) **Proposition.** Let U be a subset of a group G. The order of the stabilizer $\text{Stab}(U)$ of U for the operation of left multiplication divides the order of U.

Proof. Let H denote the stabilizer of U. Proposition (3.2) tells us that U is a union of orbits for the operation of H on G. These H-orbits are right cosets Hg. So U is a union of right cosets. Hence the order of U is a multiple of $|H|$. $\square$

Of course since the stabilizer is a subgroup of G, its order also divides $|G|$. So if $|U|$ and $|G|$ have no common factor, then $\text{Stab}(U)$ is the trivial subgroup $\{1\}$.

The operation by conjugation on subsets of G is also interesting. For example, we can partition the 15 subsets of D_3 of order 2 into orbits for conjugation. The set $\{H_1, H_2, H_3\}$ of conjugate subgroups is one orbit, and the set $\{x, x^2\}$ forms an orbit by itself. The other orbits have orders 2, 3, and 6: $15 = 1 + 2 + 3 + 3 + 6$.

For our purposes, the important thing is the orbit under conjugation of a subgroup $H \subset G$. This orbit is the set of *conjugate subgroups*

$$\{gHg^{-1} \mid g \in G\}.$$

The subgroup H is normal if and only if its orbit consists of H alone, that is, $gHg^{-1} = H$ for all $g \in G$.

The stabilizer of a subgroup H for the operation of conjugation is called the *normalizer* of H and is denoted by

(3.7) $$N(H) = \{g \in G \mid gHg^{-1} = H\}.$$

The Counting Formula reads

(3.8) $$|G| = |N(H)| \cdot |\{\text{conjugate subgroups}\}|.$$

Hence the number of conjugate subgroups is equal to the index $[G : N(H)]$.

Note that the normalizer always contains the subgroup

(3.9) $$N(H) \supset H,$$

because $hHh^{-1} = H$ when $h \in H$. So by Lagrange's Theorem, $|H|$ divides $|N(H)|$, and $|N(H)|$ divides $|G|$.

In example (3.3), the subgroups H_1, H_2, H_3 are all conjugate, and so $|N(H_i)| = 2$; hence $N(H_i) = H_i$.

The definition of the normalizer $N(H)$ shows that H is a normal subgroup of $N(H)$, and in fact $N(H)$ is the largest group containing H as a normal subgroup. In particular, $N(H) = G$ if and only if H is a normal subgroup of G.

4. THE SYLOW THEOREMS

The Sylow Theorems, which we will prove in this section, describe the subgroups of prime power order of an arbitrary finite group.

Let G be a group of order $n = |G|$, and let p be a prime number which divides n. We will use the following notation: p^e will denote the largest power of p dividing n, so that

$$n = p^e m \tag{4.1}$$

for some integer m, and p does not divide m.

(4.2) **Theorem.** *First Sylow Theorem:* There is a subgroup of G whose order is p^e.

The proofs of the Sylow Theorems are at the end of the section.

(4.3) **Corollary.** If a prime p divides the order of a finite group G, then G contains an element of order p.

For, let H be a subgroup of order p^e, and let x be an element of H different from 1. The order of x divides p^e, so it is p^r for some r in the range $0 < r \leq e$. Then $x^{p^{r-1}}$ has order p. □

Without the Sylow Theorem, this corollary is not obvious. We already know that the order of any element divides $|G|$, but we might imagine a group of order 6, for example, made up of the identity 1 and five elements of order 2. No such group exists. According to (4.3), a group of order 6 must contain an element of order 3 and an element of order 2.

(4.4) **Corollary.** There are exactly two isomorphism classes of groups of order 6. They are the classes of the cyclic group C_6 and of the dihedral group D_3.

Proof. Let x be an element of order 3 and y an element of order 2 in G. It is easily seen that the six products $x^i y^j$, $0 \leq i \leq 2$, $0 \leq j \leq 1$ are distinct elements of the group. For we can rewrite an equation $x^i y^j = x^r y^s$ in the form $x^{i-r} = y^{s-j}$. Every power of x except the identity has order 3, and every power of y except the identity has order 2. Thus $x^{i-r} = y^{s-j} = 1$, which shows that $r = i$ and $s = j$. Since G has order 6, the six elements $1, x, x^2, y, xy, x^2y$ run through the whole group. In particular, yx must be one of them. It is not possible that $yx = y$ because this would imply $x = 1$. Similarly, $yx \neq 1, x, x^2$. Therefore one of the two relations

$$yx = xy \quad \text{or} \quad yx = x^2 y$$

holds in G. Either of these relations, together with $x^3 = 1$ and $y^2 = 1$, allows us to determine the multiplication table for the group. Therefore there are at most two isomorphism classes of groups of order 6. We know two already, namely the classes of the cyclic group C_6 and of the dihedral group D_3. So they are the only ones. □

(4.5) **Definition.** Let G be a group of order $n = p^e m$, where p is a prime not dividing m and $e \geq 1$. The subgroups H of G of order p^e are called *Sylow p-subgroups* of G, or often just *Sylow subgroups.*

Thus a Sylow p-subgroup is a p-subgroup whose index in the group is not divisible by p. By Theorem (4.2), a finite group G always has a Sylow p-subgroup if p divides the order of G. The remaining Sylow Theorems (4.6) and (4.8) give more information about them.

(4.6) **Theorem.** *Second Sylow Theorem:* Let K be a subgroup of G whose order is divisible by p, and let H be a Sylow p-subgroup of G. There is a conjugate subgroup $H' = gHg^{-1}$ such that $K \cap H'$ is a Sylow subgroup of K.

(4.7) **Corollary.**

(a) If K is any subgroup of G which is a p-group, then K is contained in a Sylow p-subgroup of G.
(b) The Sylow p-subgroups of G are all conjugate.

It is clear that a conjugate of a Sylow subgroup is also a Sylow subgroup. So to obtain the first part of the corollary, we only need to note that the Sylow subgroup of a p-group K is the group K itself. So if H is a Sylow subgroup and K is a p-group, there is a conjugate H' such that $K \cap H' = K$, which is to say that H' contains K. For part (b), let K and H be Sylow subgroups. Then there is a conjugate H' of H which contains K. Since their orders are equal, $K = H'$. Thus K and H are conjugate. □

(4.8) **Theorem.** *Third Sylow Theorem:* Let $|G| = n$, and $n = p^e m$ as in (4.1). Let s be the number of Sylow p-subgroups. Then s divides m and is congruent 1 (modulo p): $s|m$, and $s = ap + 1$ for some integer $a \geq 0$.

Before proving these theorems, we will use them to determine the groups of orders 15 and 21. These examples show how powerful the Sylow Theorems are, but do not be misled. The classification of groups of order n is not easy when n has many factors. There are just too many possibilities.

(4.9) **Proposition.**

(a) Every group of order 15 is cyclic.
(b) There are two isomorphism classes of groups of order 21: the class of the cyclic group C_{21} and the class of the group G having two generators x, y which satisfy the relations $x^7 = 1$, $y^3 = 1$, $yx = x^2 y$.

Proof.

(a) Let G be a group of order 15. By (4.8) the number of its Sylow 3-subgroups divides 5 and is congruent 1 (modulo 3). The only such integer is 1. Therefore there is

one Sylow 3-subgroup H, and so it is a normal subgroup. There is one Sylow 5-subgroup K, and it is normal too, for similar reasons. Clearly, $K \cap H = \{1\}$, because the order of $K \cap H$ divides both 5 and 3. Also, KH is a subgroup of order >5, and hence $KH = G$. By (8.6) in Chapter 2, G is isomorphic to the product group $H \times K$. Thus every group of order 15 is isomorphic to a direct product of cyclic groups of orders 3 and 5. All groups of order 15 are isomorphic. Since the cyclic group C_{15} is one of them, every group of order 15 is cyclic.

(b) Let G be a group of order 21. Then Theorem (4.8) shows that the Sylow 7-subgroup K must be normal. But the possibility that there are seven conjugate Sylow 3-subgroups H is not ruled out by the theorem, and in fact this case does arise. Let x denote a generator for K, and y a generator for one of the Sylow 3-subgroups H. Then $x^7 = 1$, $y^3 = 1$, and, since K is normal, $yxy^{-1} = x^i$ for some $i < 7$.

We can restrict the possible exponents i by using the relation $y^3 = 1$. It implies that

$$x = y^3xy^{-3} = y^2x^iy^{-2} = yx^{i^2}y^{-1} = x^{i^3}.$$

Hence $i^3 \equiv 1 \pmod 7$. This means that i can take the values 1, 2, 4.

Case 1: $yxy^{-1} = x$. The group is abelian, and by (8.6) in Chapter 2 it is isomorphic to a direct product of cyclic groups of orders 3 and 7. Such a group is cyclic [Chapter 2 (8.4)].

Case 2: $yxy^{-1} = x^2$. The multiplication in G can be carried out using the rules $x^7 = 1$, $y^3 = 1$, $yx = x^2y$, to reduce every product of the elements x,y to one of the forms x^iy^j with $0 \le i < 7$ and $0 \le j < 3$. We leave the proof that this group actually exists as an exercise.

Case 3: $yxy^{-1} = x^4$. In this case, we replace y by y^2, which is also a generator for H, to reduce to the previous case: $y^2xy^{-2} = yx^4y^{-1} = x^{16} = x^2$. Thus there are two isomorphism classes of groups of order 21, as claimed. □

We will now prove the Sylow Theorems.

Proof of the First Sylow Theorem. We let $\mathcal{S}$ be the set of all subsets of G of order p^e. One of these subsets is the subgroup we are looking for, but instead of finding it directly we will show that one of these subsets has a stabilizer of order p^e. The stabilizer will be the required subgroup.

(4.10) **Lemma.** The number of subsets of order p^e in a set of $n = p^em$ elements (p not dividing m) is the binomial coefficient

$$N = \binom{n}{p^e} = \frac{n\ (n-1)\cdots(n-k)\cdots(n-p^e+1)}{p^e(p^e-1)\cdots(p^e-k)\cdots\quad 1}$$

Moreover N is not divisible by p.

Proof. It is a standard fact that the number of subsets of order p^e is this binomial coefficient. To see that N is not divisible by p, note that every time p divides a term $(n - k)$ in the numerator of N, it also divides the term $(p^e - k)$ of the denominator exactly the same number of times: If we write k in the form $k = p^i l$, where p does not divide l, then $i < e$. Therefore $(n - k)$ and $(p^e - k)$ are both divisible by p^i but not divisible by p^{i+1}. □

We decompose $\mathcal{S}$ into orbits for the operation of left multiplication, obtaining the formula

$$N = |\mathcal{S}| = \sum_{\text{orbits } O} |O|.$$

Since p does not divide N, some orbit has an order which is not divisible by p, say the orbit of the subset U. We now apply Proposition (3.6) to conclude that $|\operatorname{Stab}(U)|$ is a power of p. Since

$$(4.11) \qquad |\operatorname{Stab}(U)| \cdot |O_U| = |G| = p^e m$$

by the Counting Formula, and since $|O_U|$ is not divisible by p, it follows that $|\operatorname{Stab}(U)| = p^e$. This stabilizer is the required subgroup. □

Proof of the Second Sylow Theorem. We are given a subgroup K and a Sylow subgroup H of G, and we are to show that for some conjugate subgroup H' of H, the intersection $K \cap H'$ is a Sylow subgroup of K.

Let S denote the set of left cosets G/H. The facts that we need about this set are that G operates transitively, that is, the set forms a single orbit, and that H is the stabilizer of one of its elements, namely of $s = 1H$. So the stabilizer of as is the conjugate subgroup aHa^{-1} [see Chapter 5(6.5b)].

We restrict the operation of G to K and decompose S into K-orbits. Since H is a Sylow subgroup, the order of S is prime to p. So there is some K-orbit O whose order is prime to p. Say that O is the K-orbit of the element as. Let H' denote the stabilizer aHa^{-1} of as for the operation of G. Then the stabilizer of as for the restricted operation of K is obviously $H' \cap K$, and the index $[K:H' \cap K]$ is $|O|$, which is prime to p. Also, since it is a conjugate of H, H' is a p-group. Therefore $H' \cap K$ is a p-group. It follows that $H' \cap K$ is a Sylow subgroup of K. □

Proof of the Third Sylow Theorem. By Corollary (4.7), the Sylow subgroups of G are all conjugate to a given one, say to H. So the number of Sylow subgroups is $s = [G:N]$, where N is the normalizer of H. Since $H \subset N$, $[G:N]$ divides $[G:H] = m$. To show $s \equiv 1$ (modulo p), we decompose the set $\{H_1, \ldots, H_s\}$ of Sylow subgroups into orbits for the operation of conjugation by $H = H_1$. An orbit consists of a single group H_i if and only if H is contained in the normalizer N_i of H_i. If so, then H and H_i are both Sylow subgroups of N_i, and H_i is normal in N_i. Corollary (4.7b) shows that $H = H_i$. Therefore there is only one H-orbit of order 1, namely $\{H\}$. The other orbits have orders divisible by p because their orders divide $|H|$, by the Counting Formula. This shows that $s \equiv 1$ (modulo p). □

5. THE GROUPS OF ORDER 12

In this section, we use the Sylow Theorems to classify the groups of order 12:

(5.1) **Theorem.** There are five isomorphism classes of groups of order 12. They are represented by:

(i) the product of cyclic groups $C_3 \times C_4$;
(ii) the product of cyclic groups $C_2 \times C_2 \times C_3$;
(iii) the alternating group A_4,
(iv) the dihedral group D_6,
(v) the group generated by x, y, with relations $x^4 = 1,\ y^3 = 1,\ xy = y^2x$.

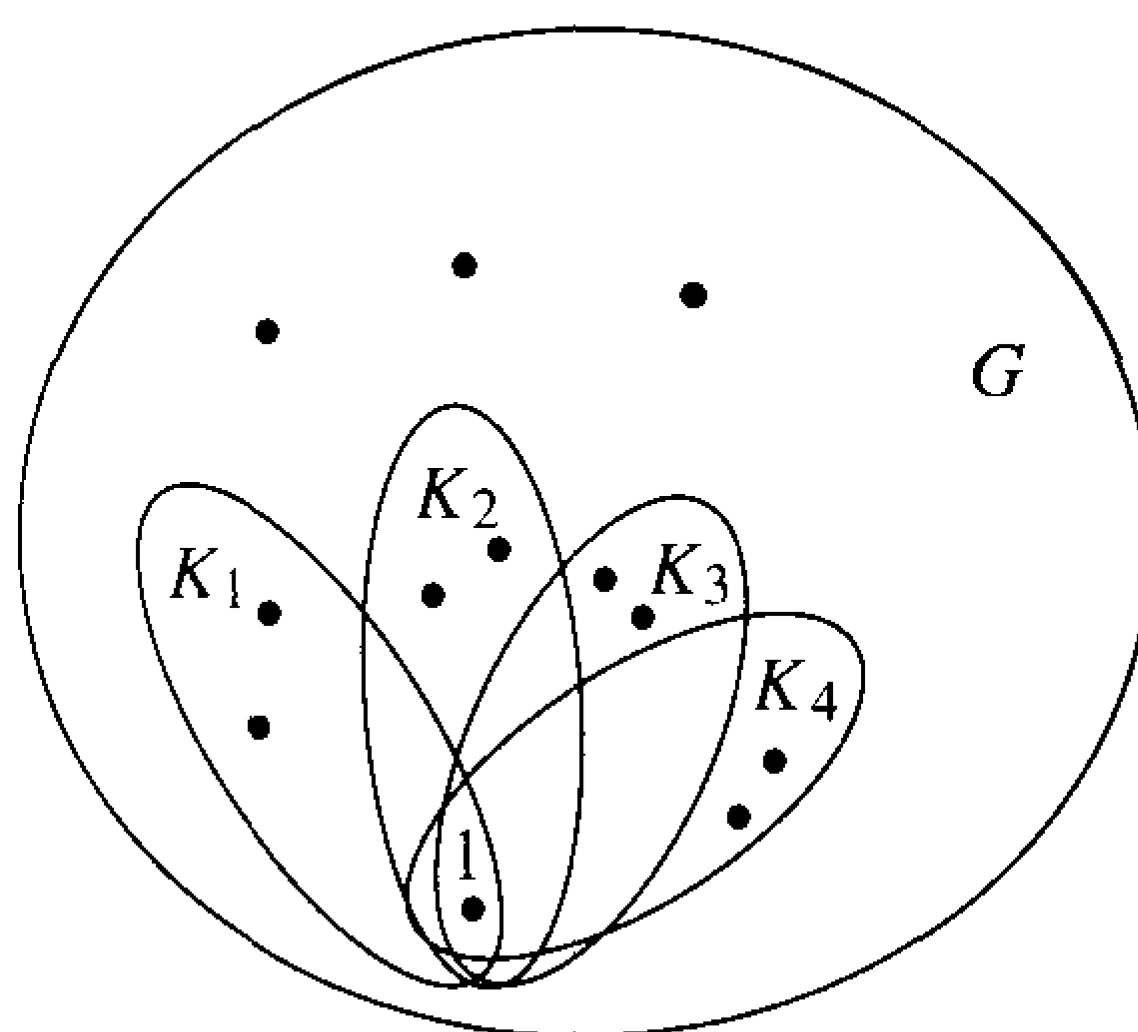
Note that $C_3 \times C_4$ is isomorphic to C_{12} and that $C_2 \times C_2 \times C_3$ is isomorphic to $C_2 \times C_6$ (see [Chapter 2 (8.4)]).

Proof. Let G be a group of order 12. Denote by H a Sylow 2-subgroup of G, which has order 4, and by K a Sylow 3-subgroup, of order 3. It follows from Theorem (4.8) that the number of Sylow 2-subgroups is either 1 or 3, and that the number of Sylow 3-subgroups is 1 or 4. Also, H is a group of order 4 and is therefore either a cyclic group or the Klein four group V, a product of two cyclic groups of order 2:

(5.2) $$H \approx C_4 \quad \text{or} \quad H \approx V.$$

(5.3) **Lemma.** At least one of the two subgroups H,K is normal.

Proof. Suppose that K is not normal. Then K has four conjugate subgroups $K = K_1, \ldots, K_4$. Since $|K_i| = 3$, the intersection of any two of these groups must be the identity. Counting elements shows that there are only three elements of G which are not in any of the groups K_i.



Any Sylow 2-subgroup H has order 4, and $H \cap K_i = \{1\}$. Therefore it consists of these three elements and 1. This describes H for us and shows that there is only one Sylow 2-subgroup. Thus H is normal. □

Since $H \cap K = \{1\}$, every element of HK has a unique expression as a product hk [Chapter 2 (8.6)], and since $|G| = 12$, $HK = G$. If H is normal, then K operates on H by conjugation, and we will show that this operation, together with the structure of H and K, determines the structure of G. Similarly, if K is normal then H operates on K, and this operation determines G.

Case 1: H and K are both normal. Then by (8.6) in Chapter 2, G is isomorphic to the product group $H \times K$. By (5.2) there are two possibilities:

(5.4) $$G \approx C_4 \times C_3 \quad \text{or} \quad G \approx V \times C_3.$$

These are the abelian groups of order 12.

Case 2: H is normal but K is not. So there are four conjugate Sylow 3-subgroups $\{K_1,\ldots,K_4\}$, and G operates by conjugation on this set S of four subgroups. This operation determines a permutation representation

(5.5) $$G \xrightarrow{\varphi} S_4.$$

Let us show that φ maps G isomorphically to the alternating group A_4 in this case.

The stabilizer of K_i for the operation of conjugation is the normalizer $N(K_i)$, which contains K_i. The Counting Formula shows that $|N(K_i)| = 3$, and hence that $N(K_i) = K_i$. Since the only element common to the subgroups K_i is the identity element, only the identity stabilizes all of these subgroups. Thus φ is injective and G is isomorphic to its image in S_4.

Since G has four subgroups of order 3, it contains eight elements of order 3, and these elements certainly generate the group. If x has order 3, then $\varphi(x)$ is a permutation of order 3 in S_4. The permutations of order 3 are even. Therefore $\operatorname{im}\varphi \subset A_4$. Since $|G| = |A_4|$, the two groups are equal.

As a corollary, we note that if H is normal and K is not, then H is the Klein four group V, because the Sylow 2-subgroup of A_4 is V.

Case 3: K is normal, but H is not. In this case H operates on K by conjugation, and conjugation by an element of H is an automorphism of K. We let y be a generator for the cyclic group K: $y^3 = 1$. There are only two automorphisms of K—the identity and the automorphism which interchanges y and y^2.

Suppose that H is cyclic of order 4, and let x generate H: $x^4 = 1$. Then since G is not abelian, $xy \neq yx$, and so conjugation by x is not the trivial automorphism of K. Hence $xyx^{-1} = y^2$. The Todd–Coxeter Algorithm (see Section 9) is one way to show that these relations define a group of order 12:

(5.6) $$x^4 = 1\,,\ y^3 = 1\,,\ xyx^{-1} = y^2.$$

The last possibility is that H is isomorphic to the Klein four group. Since there are only two automorphisms of K, there is an element $w \in H$ besides the identity which operates trivially: $wyw^{-1} = y$. Since G is not abelian, there is also an element v which operates nontrivially: $vyv^{-1} = y^2$. Then the elements of H are $\{1,v,w,vw\}$, and the relations $v^2 = w^2 = 1$, and $vw = wv$ hold in H. The element $x = wy$ has

order 6, and $vxv^{-1} = vwyv^{-1} = wy^2 = y^2w = x^{-1}$. The relations $x^6 = 1$, $v^2 = 1$, $vxv^{-1} = x^{-1}$ define the group D_6, so G is dihedral in this case. □

6. COMPUTATION IN THE SYMMETRIC GROUP

We want to bring up two points about calculation with permutations. The first concerns the order of multiplication. To have a uniform convention, we have used the functional notation $p(x)$ for all our maps p, including permutations. This has the consequence that a product pq must be interpreted as the composed operation $p \circ q$, that is, "first apply q, then p." When multiplying permutations, it is more usual to read pq as "first apply p, then q." We will use this second convention here. A compatible notation for the operation of a permutation p on an index $\mathbf{i}$ requires writing the permutation on the right side of the index:

$$(\mathbf{i})p.$$

Applying first p and then q to an index $\mathbf{i}$, we get $((\mathbf{i})p)q = (\mathbf{i})pq$, as desired. Actually, this notation looks funny to me. We will usually drop the parentheses:

$$(\mathbf{i})p = \mathbf{i}p.$$

What is important is that p must appear on the right.

To make our convention for multiplication compatible with matrix multiplication, we must replace the matrix P associated to a permutation p in Chapter 1 (4.6) by its transpose P^t, and use it to multiply on the right on a row vector.

The second point is that it is not convenient to compute with permutation matrices, because the matrices are large in relation to the amount of information they contain. A better notation is needed. One way to describe a permutation is by means of a table. We can consider the configuration

$$p = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 6 & 8 & 3 & 5 & 2 & 1 & 7 \end{bmatrix} \tag{6.1}$$

as a notation for the permutation defined by

$$\mathbf{1}p = \mathbf{4},\ \mathbf{2}p = \mathbf{6}, \ldots .$$

It is easy to compute products using this notation. If for example

$$q = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 4 & 6 & 8 & 1 & 3 & 5 & 7 \end{bmatrix},$$

then we can evaluate pq (first p, then q) by reading the two tables in succession:

$$pq = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 8 & 3 & 7 & 6 & 1 & 4 & 2 & 5 \end{bmatrix}.$$

Table (6.1) still requires a lot of writing, and of course the top row is always the same. It could, in principle, be left off, to reduce the amount of writing by half,

but this would make it hard to find our place in the bottom row if we were permuting, say, 18 digits.

Another notation, called *cycle notation,* is commonly used. It describes a permutation of n elements by at most n symbols and is based on the partition of the indices into orbits for the operation of a permutation. Let p be a permutation, and let H be the cyclic subgroup generated by p. We decompose the set $\{\mathbf{1},\dots,\mathbf{n}\}$ into H-orbits and refer to these orbits as the *p-orbits*. The p-orbits form a partition of the set of indices, called the *cycle decomposition* associated to the permutation p.

If an index $\mathbf{i}$ is in an orbit of k elements, the elements of the orbit will be

$$O = \{\mathbf{i}, \mathbf{i}p, \mathbf{i}p^2, \dots, \mathbf{i}p^{k-1}\}.$$

Let us denote $\mathbf{i}p^r$ by $\mathbf{i}_r$, so that $O = \{\mathbf{i}_0, \mathbf{i}_1, \dots, \mathbf{i}_{k-1}\}$. Then p operates on this orbit as

(6.2)

$$\mathbf{i}_0 \to \mathbf{i}_1 \to \mathbf{i}_2 \to \cdots \to \mathbf{i}_{k-1} \to \mathbf{i}_0$$

A permutation which operates in this way on a subset $\{\mathbf{i}_0, \mathbf{i}_1, \dots, \mathbf{i}_{k-1}\}$ of the indices and leaves the remaining indices fixed is called a *cyclic permutation*. Thus

(6.3)

$$\sigma = \mathbf{1} \to \mathbf{4} \to \mathbf{3} \to \mathbf{8} \to \mathbf{7} \to \mathbf{1}$$

defines a cyclic permutation of order 5 of $\{\mathbf{1},\dots,\mathbf{8}\}$, it being understood that the indices $\mathbf{2}, \mathbf{5}, \mathbf{6}$ which are not mentioned are left fixed—each forms a σ-orbit of one element. When we speak of *the indices on which a permutation operates,* we will mean the ones which are not fixed: $\mathbf{1}, \mathbf{3}, \mathbf{4}, \mathbf{7}, \mathbf{8}$ in this case.

Another cyclic permutation of $\{\mathbf{1},\dots,\mathbf{8}\}$ is

(6.4)

$$\tau = \mathbf{2} \to \mathbf{6} \to \mathbf{2}$$

Such a cyclic permutation of order 2 is called a *transposition*. A transposition is a permutation which operates on two indices.

Our permutation p (6.1) is not cyclic because there are three p-orbits:

$$p: \quad \underbrace{\mathbf{1} \to \mathbf{4} \to \mathbf{3} \to \mathbf{8} \to \mathbf{7} \to \mathbf{1}}_{\sigma} \qquad \mathbf{5} \to \mathbf{5} \qquad \underbrace{\mathbf{2} \to \mathbf{6} \to \mathbf{2}}_{\tau}$$

It is clear that

$$p = \sigma\tau = \tau\sigma,$$

where $\sigma\tau$ denotes the product permutation.

(6.5) **Proposition.** Let σ, τ be permutations which operate on disjoint sets of indices. Then $\sigma\tau = \tau\sigma$.

Proof. If neither σ nor τ operates on an index $\mathbf{i}$, then $\mathbf{i}\sigma\tau = \mathbf{i}\tau\sigma = \mathbf{i}$. If σ sends $\mathbf{i}$ to $\mathbf{j} \neq \mathbf{i}$, then τ fixes both $\mathbf{i}$ and $\mathbf{j}$. In that case, $\mathbf{i}\sigma\tau = \mathbf{j}\tau = \mathbf{j}$ and $\mathbf{i}\tau\sigma = \mathbf{i}\sigma = \mathbf{j}$ too. The case that τ operates on $\mathbf{i}$ is the same. □

Note, however, that when we multiply permutations which operate on overlapping sets of indices, the operations need not commute. The symmetric group S_n is not a commutative group if $n > 2$. For example, if τ' is the transposition which interchanges 3 and 6 and if σ is as above, then $\sigma\tau' \neq \tau'\sigma$.

(6.6) **Proposition.** Every permutation p not the identity is a product of cyclic permutations which operate on disjoint sets of indices: $p = \sigma_1\sigma_2 \cdots \sigma_k$, and these cyclic permutations σ_r are uniquely determined by p.

Proof. We know that p operates as a cyclic permutation when restricted to a single orbit. For each p-orbit, we may define a cyclic permutation σ_r which permutes that orbit in the same way that p does and which fixes the other indices. Clearly, p is the product of these cyclic permutations. Conversely, let p be written as a product $\sigma_1\sigma_2 \cdots \sigma_k$ of cyclic permutations operating on distinct sets $O_1, \ldots, O_k$ of indices. According to Proposition (6.5), the order does not matter. Note that $\sigma_2, \ldots, \sigma_k$ fix the elements of O_1; hence p and σ_1 act in the same way on O_1. Therefore O_1 is a p-orbit. The same is true for the other cyclic permutations. Thus $O_1, \ldots, O_k$ are the p-orbits which contain more than one element, and the permutations σ_i are those constructed at the start of the proof. □

A *cycle notation* for the cyclic permutation (6.2) is

$$(\mathbf{i}_0\mathbf{i}_1 \cdots \mathbf{i}_{k-1}). \tag{6.7}$$

Thus our particular permutation σ has the cycle notation $(\mathbf{14387})$. The notation is not completely determined by the permutation, because we can start the list with any of the indices $\mathbf{i}_0, \ldots, \mathbf{i}_{k-1}$. There are five equivalent notations for σ:

$$\sigma = (\mathbf{43871}) = (\mathbf{38714}) = \cdots.$$

Any one of these notations may be used.

A *cycle notation* for an arbitrary permutation p is obtained by writing the permutation as a product of cyclic permutations which operate on disjoint indices, and then writing the cycle notations for each of these permutations in succession. The order is irrelevant. Thus two of the possible cycle notations for the permutation p described above are

$$(\mathbf{14387})(\mathbf{26}) \quad \text{and} \quad (\mathbf{62})(\mathbf{87143}).$$

If we wish, we can include the "one-cycle" $(\mathbf{5})$, to represent the fixed element $\mathbf{5}$, thereby presenting all the indices in the list. But this is not customary.

With this notation, every permutation can be denoted by a string of at most n integers, suitably bracketed. Products can still be described by juxtaposition. A cycle notation for the permutation q considered above is $q = (\mathbf{1\,2\,4\,8\,7\,5})(\mathbf{3\,6})$. Thus

$$pq = \overset{\sigma}{(\mathbf{14387})}\overset{\tau}{(\mathbf{26})}\overset{\sigma'}{(\mathbf{124875})}\overset{\tau'}{(\mathbf{36})} = \sigma\tau\sigma'\tau'. \tag{6.8}$$

This string of cycles represents the permutation pq. To evaluate the product on an index, the index is followed through the four factors:

$$\mathbf{1} \overset{\sigma}{\rightsquigarrow} \mathbf{4} \overset{\tau}{\rightsquigarrow} \mathbf{4} \overset{\sigma'}{\rightsquigarrow} \mathbf{8} \overset{\tau'}{\rightsquigarrow} \mathbf{8}, \text{ and so on.}$$

However, (6.8) does not exhibit the decomposition of pq into disjoint cycles, because indices appear more than once. Computation of the permutation as above leads to the cycle decomposition

$$pq = (\mathbf{185})(\mathbf{237})(\mathbf{46}) = \mathbf{1}\to\mathbf{8}\to\mathbf{5}\to\mathbf{1} \quad \mathbf{2}\to\mathbf{3}\to\mathbf{7}\to\mathbf{2} \quad \mathbf{4}\leftrightarrow\mathbf{6}.$$

When the computation is finished, every index occurs at most once.

For another sample, let $\rho = (\mathbf{5\,4\,8})$. Then

$$\begin{aligned} \sigma\rho &= (\mathbf{14387})(\mathbf{548}) = (\mathbf{187})(\mathbf{354}) \\ \rho\sigma &= (\mathbf{548})(\mathbf{14387}) = (\mathbf{147})(\mathbf{385}). \end{aligned} \tag{6.9}$$

Now let us compute the conjugate of a permutation p. Since p is a product of disjoint cycles, it will be enough to describe the conjugate $q^{-1}\sigma q$ of a cyclic permutation σ, say the permutation $(\mathbf{i}_1 \cdots \mathbf{i}_k)$. (The fact that we have switched the order of multiplication makes the expression for conjugation by q^{-1} a little nicer than that for conjugation by q.)

(6.10) **Proposition.**

(a) Let σ denote the cyclic permutation $(\mathbf{i}_1\mathbf{i}_2 \cdots \mathbf{i}_k)$, and let q be any permutation. Denote the index $\mathbf{i}_r q$ by $\mathbf{j}_r$. Then the conjugate permutation $q^{-1}\sigma q$ is the cyclic permutation $(\mathbf{j}_1\mathbf{j}_2 \cdots \mathbf{j}_k)$.

(b) If an arbitrary permutation p is written as a product of disjoint cycles σ, then $q^{-1}pq$ is the product of the disjoint cycles $q^{-1}\sigma q$.

(c) Two permutations p, p' are conjugate elements of the symmetric group if and only if their cycle decompositions have the same orders.

Proof. The proof of (a) is the following computation:

$$\mathbf{j}_r q^{-1}\sigma q = \mathbf{i}_r \sigma q = \mathbf{i}_{r+1} q = \mathbf{j}_{r+1},$$

in which the indices are to be read modulo k. Part (b) follows easily. Also, the fact that conjugate permutations have cycle decompositions with the same orders follows from (b). Conversely, suppose that p and p' have cycle decompositions of the same

orders. Say that $p = (\mathbf{i}_1 \cdots \mathbf{i}_r)(\mathbf{i}_1' \cdots \mathbf{i}_s') \cdots$ and $p' = (\mathbf{j}_1 \cdots \mathbf{j}_r)(\mathbf{j}_1' \cdots \mathbf{j}_s') \cdots$. Define q to be the permutation sending $\mathbf{i}_\nu \rightsquigarrow \mathbf{j}_\nu$, $\mathbf{i}_\nu' \rightsquigarrow \mathbf{j}_\nu'$, and so on. Then $p' = q^{-1}pq$. □

Let us determine the Class Equation for the symmetric group S_4 as an example. This group contains six transpositions

$$(\mathbf{1\,2}),\quad (\mathbf{1\,3}),\quad (\mathbf{1\,4}),\quad (\mathbf{2\,3}),\quad (\mathbf{2\,4}),\quad (\mathbf{3\,4}),$$

three products of disjoint transpositions

$$(\mathbf{1\,2})(\mathbf{3\,4}),\quad (\mathbf{1\,3})(\mathbf{2\,4}),\quad (\mathbf{1\,4})(\mathbf{2\,3}),$$

eight 3-cycles, and six 4-cycles. By Proposition (6.10), each of these sets forms one conjugacy class. So the Class Equation of S_4 is

$$24 = 1 + 3 + 6 + 6 + 8.$$

We will now describe the subgroups G of the symmetric group S_p whose order is divisible by p and whose Sylow p-subgroup is normal. We assume that p is a prime integer. Since p divides $p! = |S_p|$ only once, it also divides $|G|$ once, and so the Sylow p-subgroup of G is a cyclic group.

It turns out that such subgroups have a very nice description in terms of the finite field $\mathbb{F}_p$. To obtain it, we use the elements $\{\mathbf{0}, \mathbf{1}, \cdots, \mathbf{p-1}\}$ of the finite field as the indices. Certain permutations of this set are given by the field operations themselves. Namely, we have the operations (*add a*) and (*multiply by c*) for any given $a, c \in \mathbb{F}_p$, $c \neq 0$. They are invertible operations and hence permutations of $\mathbb{F}_p$, so they represent elements of the symmetric group. For example, (*add* 1) is the p-cycle

$$(6.11) \qquad (\textit{add } 1) = (\mathbf{0\,1\,2} \cdots (\mathbf{p-1})).$$

The operator (*multiply by c*) always fixes the index $\mathbf{0}$, but its cycle decomposition depends on the order of the element c in $\mathbb{F}_p^\times$. For example,

$$(6.12) \qquad \begin{aligned} (\textit{multiply by } 2) &= (\mathbf{1\,2\,4\,3}) && \text{if } p = 5 \\ &= (\mathbf{1\,2\,4})(\mathbf{3\,6\,5}) && \text{if } p = 7. \end{aligned}$$

Combining the operations of addition and multiplication gives us all operators on $\mathbb{F}_p$ of the form

$$(6.13) \qquad x \rightsquigarrow cx + a.$$

The set of these operators forms a subgroup G of order $p(p-1)$ of the symmetric group.

The group of operators (6.13) has a nice matrix representation, as the set of 2×2 matrices with entries in the field $\mathbb{F}_p$, of the form

$$(6.14) \qquad \begin{bmatrix} 1 & a \\ & c \end{bmatrix}.$$

This matrix operates by right multiplication on the vector $(1,x)$, sending it to $(1, cx + a)$. So we can recover the operation of G on $\mathbb{F}_p$ from right multiplication by the corresponding matrix. (We use right multiplication because of our chosen order of operations.) The operations (add a) and (multiply by c) are represented by the elementary matrices

$$\begin{bmatrix} 1 & a \\ & 1 \end{bmatrix}, \begin{bmatrix} 1 & \\ & c \end{bmatrix}.$$

(6.15) **Theorem.** Let p be a prime, and let H be a subgroup of the symmetric group S_p whose order is divisible by p. If the Sylow p-subgroup of H is normal, then, with suitable labeling of the indices, H is contained in the group of operators of the form (6.13).

For example, the dihedral group D_p operates faithfully on the vertices of a regular p-gon, and so it is realized as a subgroup of the symmetric group S_p. It is the subgroup of (6.14) consisting of the matrices in which $c = \pm 1$.

Proof of the theorem. The only elements of order p of S_p are the p-cycles. So H contains a p-cycle, say σ. We may relabel indices so that σ becomes the standard p-cycle (*add* 1) $= (\mathbf{0}\,\mathbf{1}\cdots(\mathbf{p}-\mathbf{1}))$. Then this permutation generates the Sylow p-subgroup of H.

Let τ_1 be another element of H. We have to show that τ_1 corresponds to an operator of the form (6.13). Say that τ_1 sends the index $\mathbf{0}$ to $\mathbf{i}$. Since σ^i also sends $\mathbf{0}$ to $\mathbf{i}$, the product $\tau = \sigma^{-i}\tau_1$ fixes $\mathbf{0}$. It suffices to show that τ has the form (6.13), and to do so, we will show that τ is one of the operators (*multiply by c*).

By hypothesis, $K = \{1, \sigma, \ldots, \sigma^{p-1}\}$ is a normal subgroup of H. Therefore

$$\tau^{-1}\sigma\tau = \sigma^k \tag{6.16}$$

for some k between 1 and $p-1$. We now determine τ by computing both sides of this equation. By Proposition (6.10), the left side is the p-cycle $\tau^{-1}\sigma\tau = (\mathbf{0}\tau\,\mathbf{1}\tau\ldots(\mathbf{p}-\mathbf{1})\tau)$, while direct computation of the right side gives $\sigma^k = (\mathbf{0k}\;\mathbf{2k}\ldots(\mathbf{p}-\mathbf{1})\mathbf{k})$:

$$(\mathbf{0}\tau\,\mathbf{1}\tau\ldots(\mathbf{p}-\mathbf{1})\tau) = (\mathbf{0}\;\mathbf{k}\,\mathbf{2k}\ldots(\mathbf{p}-\mathbf{1})\mathbf{k}).$$

We must be careful in interpreting the equality of these two cycles, because the cycle notation is not unique. We need to know that the first index on the left is the same as the first index on the right. Otherwise we will have to identify equal indices in the two cycles and begin with them. That is why we normalized at the start, to have $\mathbf{0}\tau = \mathbf{0}$. Knowing that fact, the two lists are the same, and we conclude that

$$\mathbf{1}\tau = \mathbf{k}, \quad \mathbf{2}\tau = \mathbf{2k}, \quad \ldots .$$

This is the operator (*multiply by k*), as claimed. □

We now return for a moment to the question of order of operations. If we wish to use the notation $p(\mathbf{i})$ for permutations in this section, as we do for functions else-

where, we must modify our way of computing with cycles in order to take this into account. The most systematic way to proceed is to read *everything*, including cycles, from right to left. In other words, we should read the cycle (**14387**) as

$$\mathbf{1} \leftsquigarrow \mathbf{4} \leftsquigarrow \mathbf{3} \leftsquigarrow \mathbf{8} \leftsquigarrow \mathbf{7} \leftsquigarrow \mathbf{1}.$$

This is the inverse of the permutation (6.3). We can then interpret the product (**14387**)(**548**) as composition: "First apply (**548**), then (**14387**)." Computation of this product gives

$$\mathbf{1} \leftsquigarrow \mathbf{8} \leftsquigarrow \mathbf{7} \leftsquigarrow \mathbf{1}, \quad \mathbf{3} \leftsquigarrow \mathbf{5} \leftsquigarrow \mathbf{4} \leftsquigarrow \mathbf{3},$$

which we would write as (**187**)(**354**). Notice that this is the same string of symbols as we obtained in (6.9). Miraculously, reading everything backward gives the same answer when we multiply permutations. But of course, the notation (**187**)(**354**) now stands for the inverse of the permutation (6.9). The fact that the notations multiply consistently in our two ways of reading permutations mitigates the crime we have committed in switching from left to right.

7. THE FREE GROUP

We have seen a few groups, such as the symmetric group S_3, the dihedral groups D_n, and the group M of rigid motions of the plane, in which one can compute easily using a list of generators and a list of relations for manipulating them. The rest of this chapter is devoted to the formal background for such methods. In this section, we consider groups which have a set of generators satisfying *no* relations other than ones [such as $x(yz) = (xy)z$] which are implied by the group axioms. A set S of elements of a group which satisfy no relations except those implied by the axioms is called *free*, and a group which has a free set of generators is called a *free group*. We will now describe the free groups.

We start with an arbitrary set S of symbols, say $S = \{a, b, c, \ldots\}$, which may be finite or infinite, and define a *word* to be a finite string of symbols from S, in which repetition is allowed. For instance a, aa, ba, and $aaba$ are words. Two words can be composed by juxtaposition:

$$aa, ba \rightsquigarrow aaba;$$

in this way the set W of all words has an associative law of composition. Moreover, the "empty word" can be introduced as an identity element for this law. We will need a symbol to denote the empty word; let us use 1. The set W is called the *free semigroup* on the set of symbols S. Unfortunately it is not a group because inverses are lacking, and the introduction of inverses complicates things.

Let S' be the set consisting of the symbols in S and also of symbols a^{-1} for every $a \in S$:

$$S' = \{a, a^{-1}, b, b^{-1}, c, c^{-1}, \ldots\}. \tag{7.1}$$

Let W' be the set of words made using the symbols S'. If a word $w \in W'$ looks

like

$$\cdots xx^{-1} \cdots \quad \text{or} \quad \cdots x^{-1}x \cdots$$

for some $x \in S$, then we can agree to *cancel* the two symbols x, x^{-1} and reduce the length of the word. The word will be called *reduced* if no such cancellation can be made. Starting with any word w, we can perform a finite sequence of cancellations and must eventually get a reduced word w_0, possibly the empty word 1. We call this word w_0 a *reduced form* of w.

Now there is often more than one way to proceed with cancellation. For instance, starting with $w = babb^{-1}a^{-1}c^{-1}ca$, we can proceed in several ways, such as

$$\begin{array}{cc}
\underline{b}a\not{b}\not{b}^{-1}a^{-1}c^{-1}c\underline{a} & \underline{ba}bb^{-1}a^{-1}\not{c}^{-1}\not{c}a \\
\downarrow & \downarrow \\
\underline{b}\not{a}\not{a}^{-1}c^{-1}c\underline{a} & \underline{ba}bb^{-1}\not{a}^{-1}\not{a} \\
\downarrow & \downarrow \\
\underline{b}\not{c}^{-1}\not{c}\underline{a} & \underline{ba}\not{b}\not{b}^{-1} \\
\downarrow & \downarrow \\
ba & ba.
\end{array}$$

The same reduced word is obtained at the end, though the letters come from different places in the original word. (The letters which remain at the end have been underlined.) This is the general situation.

(7.2) **Proposition.** There is only one reduced form of a given word w.

Proof. We use induction on the length of w. If w is reduced, there is nothing to show. If not, there must be some pair of letters which can be cancelled, say the underlined pair

$$w = \cdots \underline{xx^{-1}} \cdots.$$

(Let us allow x to denote any element of S', with the obvious convention that if $x = a^{-1}$ then $x^{-1} = a$.) If we show that we can obtain every reduced form w_0 of w by cancelling the pair $\underline{xx^{-1}}$ first, then the proposition will follow by induction on the shorter word $\cdots \underline{\not{x}\not{x}^{-1}} \cdots$ thus obtained.

Let w_0 be a reduced form of w. We know that w_0 is obtained from w by some sequence of cancellations. The first case is that our pair $\underline{xx^{-1}}$ is cancelled at some step in this sequence. Then we might as well rearrange the operations and cancel $\underline{xx^{-1}}$ first. So this case is settled. On the other hand, the pair $\underline{xx^{-1}}$ can not remain in w_0, since w_0 is reduced. Therefore at least one of the two symbols must be cancelled at some time. If the pair itself is not cancelled, then the first cancellation involving the pair must look like

$$\cdots \not{x}^{-1}\underline{\not{x}x^{-1}} \cdots \quad \text{or} \quad \cdots \underline{x\not{x}^{-1}}\not{x} \cdots.$$

Notice that the word obtained by this cancellation is the same as that obtained by

cancelling the original pair $\underline{xx^{-1}}$. So we may cancel the original pair at this stage instead. Then we are back in the first case, and the proposition is proved. □

Now we call two words w, w' in W' *equivalent*, and we write $w \sim w'$, if they have the same reduced form. This is an equivalence relation.

(7.3) **Proposition.** The product of equivalent words is equivalent: If $w \sim w'$ and $v \sim v'$, then $wv \sim w'v'$.

Proof. To obtain the reduced word equivalent to the product wv, we can first cancel as much as possible in w and in v, to reduce w to w_0 and v to v_0. Then wv is reduced to w_0v_0. Now we continue cancelling in w_0v_0 if possible. Since $w' \sim w$ and $v' \sim v$, the same process, applied to $w'v'$, passes through w_0v_0 too, and hence it leads to the same reduced word. □

It follows from this proposition that equivalence classes of words may be multiplied, that is, that there is a well-defined law of composition on the set of equivalence classes of words.

(7.4) **Proposition.** Let F denote the set of equivalence classes of words in W'. Then F is a group with the law of composition induced from W'.

Proof. The facts that multiplication is associative and that the class of the empty word 1 is an identity follow from the corresponding facts in W'. It remains to check that all elements of F are invertible. But clearly, if $w = xy \cdots z$ then the class of $z^{-1} \cdots y^{-1}x^{-1}$ is the inverse of the class of w. □

(7.5) **Definition.** The group F of equivalence classes of words is called the *free group* on the set S.

So an element of the free group F corresponds to exactly one reduced word in W', by Proposition (7.2). To multiply reduced words, combine and cancel:

$$(abc^{-1})(cb) \rightsquigarrow abc^{-1}cb = abb.$$

One can also introduce power notation for reduced words: $aaab^{-1}b^{-1} = a^3b^{-2}$.

The free group on the set $S = \{a\}$ consisting of one element is the same as the set of all powers of a: $F = \{a^n\}$. It is an infinite cyclic group. In contrast, the free group on a set $S = \{a, b\}$ of two elements is very complicated.

8. GENERATORS AND RELATIONS

Having described free groups, we now consider the more likely case that a set of generators of a group is not free—that there are some nontrivial relations among them. Our discussion is based on the mapping properties of the free group and of quotient groups.

(8.1) **Proposition.** *Mapping property of the free group:* Let F be the free group on a set $S = \{a, b, \ldots\}$, and let G be a group. Every map of sets $f: S \longrightarrow G$ extends in a unique way to a group homomorphism $\varphi: F \longrightarrow G$. If we denote the image $f(x)$ of an element $x \in S$ by $\tilde{x}$, then φ sends a word in $S' = \{a, a^{-1}, b, b^{-1}, \ldots\}$ to the corresponding product of the elements $\{\tilde{a}, \tilde{a}^{-1}, \tilde{b}, \tilde{b}^{-1}, \ldots\}$ in G.

Proof. This rule does define a map on the set of words in S'. We must show that equivalent words are sent to the same product in G. But since cancellation in a word will not change the corresponding product in G, this is clear. Also, since multiplication in F is defined by juxtaposition, the map φ thus defined is a homomorphism. It is the only way to extend f to a homomorphism. □

If S is any subset of a group G, the mapping property defines a homomorphism $\varphi: F \longrightarrow G$ from the free group on S to G. This reflects the fact that the elements of S satisfy no relations in F except those implied by the group axioms, and explains the reason for the adjective *free*.

A family S of elements is said to *generate* a group G if the map φ from the free group on S to G is surjective. This is the same as saying that every element of G is a product of some string of elements of S', so it agrees with the terminology introduced in Section 2 of Chapter 2. In any case, whether or not S generates G, the image of the homomorphism φ of Proposition (8.1) is a subgroup called the *subgroup generated by S*. This subgroup consists precisely of all products of elements of S'.

Assume that S generates G. The elements of S are then called *generators*. Since φ is a surjective homomorphism, the First Isomorphism Theorem [Chapter 2 (10.9)] tells us that G is isomorphic to the quotient group F/N, where $N = \ker \varphi$. The elements of N are called *relations* among the generators. They are equivalence classes of words w with the property that the corresponding product in G is 1:

$$\varphi(w) = 1 \quad \text{or} \quad w = 1 \textit{ in } G.$$

In the special case that $N = \{1\}$, φ is an isomorphism. In this case G is called a free group too.

If we know a set of generators and also all the relations, then we can compute in the isomorphic group F/N and hence in our group G. But the subgroup N will be infinite unless G is free, so we can't list all its elements. Rather, a set of words

$$R = \{r_1, r_2, \ldots\}$$

is called a set of *defining relations* for G if $R \subset N$ and if N is the *smallest normal subgroup containing* R. This means that N is generated by the subset consisting of all the words in R and also all their conjugates.

It might seem more systematic to require the defining relations to be generators for the group N. But remember that the kernel of the homomorphism $F \longrightarrow G$ defined by a set of generators is always a normal subgroup, so there is no need to make the list of defining relations longer. If we know that some relation $r = 1$ holds in G, then we can conclude that $grg^{-1} = 1$ holds in G too, simply by multiplying both sides of the equation on the left and right by g and g^{-1}.

We already know a few examples of generators and relations, such as the dihedral group D_n [Chapter 5 (3.6), (3.7)]. It is generated by the two elements x, y, with relations

$$x^n = 1, \quad y^2 = 1, \quad xyxy = 1. \tag{8.2}$$

(8.3) **Proposition.** The elements $x^n, y^2, xyxy$ form a set of defining relations for the dihedral group.

This proposition is essentially what was checked in Chapter 5 (3.6). But to prove it formally, and to work freely with the concept of generators and relations, we will need what is called the mapping property of quotient groups. It is a generalization of the First Isomorphism Theorem:

(8.4) **Proposition.** *Mapping property of quotient groups:* Let N be a normal subgroup of G, let $\overline{G} = G/N$, and let π be the canonical map $G \longrightarrow \overline{G}$ defined by $\pi(a) = \overline{a} = aN$. Let φ: $G \longrightarrow G'$ be a homomorphism whose kernel contains N. There is a unique homomorphism $\overline{\varphi}$: $\overline{G} \longrightarrow G'$ such that $\overline{\varphi}\pi = \varphi$:

$$\begin{array}{ccc} G & \xrightarrow{\varphi} & G' \\ & {\scriptstyle\pi}\searrow \quad \nearrow{\scriptstyle\overline{\varphi}} & \\ & \overline{G} & \end{array}$$

This map is defined by the rule $\overline{\varphi}(\overline{a}) = \varphi(a)$.

Proof. To define a map $\overline{\varphi}$: $\overline{G} \longrightarrow G'$, we must define $\overline{\varphi}(\alpha)$ for every element α of $\overline{G}$. To do this, we represent α by an element $a \in G$, choosing a so that $\alpha = \pi(a)$. In the bar notation, this means that $\alpha = \overline{a}$. Now since we want our map $\overline{\varphi}$ to satisfy the relation $\overline{\varphi}(\pi(a)) = \varphi(a)$, there is no choice but to define $\overline{\varphi}$ by the rule $\overline{\varphi}(\alpha) = \varphi(a)$, as asserted in the proposition. To show that this is permissible, we must show that the value we obtained for $\overline{\varphi}(\alpha)$, namely $\varphi(a)$, depends only on α and not on our choice of the representative a. This is often referred to as showing that our map is "well-defined."

Let a and a' be two elements of G such that $\overline{a} = \overline{a}' = \alpha$. The equality $\overline{a} = \overline{a}'$ means that $aN = a'N$, or [Chapter 2 (5.13)] that $a' \in aN$. So $a' = an$ for some $n \in N$. Since $N \subset \ker \varphi$ by hypothesis, $\varphi(n) = 1$. Thus $\varphi(a') = \varphi(a)\varphi(n) = \varphi(a)$, as required.

Finally, the map $\overline{\varphi}$ is a homomorphism because $\overline{\varphi}(\overline{a})\overline{\varphi}(\overline{b}) = \varphi(a)\varphi(b) = \varphi(ab) = \overline{\varphi}(\overline{ab})$. □

Proof of Proposition (8.3). We showed in Chapter 5 (3.6) that D_n is generated by elements x, y which satisfy (8.2). Therefore there is a surjective map φ: $F \longrightarrow D_n$ from the free group on x, y to D_n, and $R = \{x^n, y^2, xyxy\}$ is contained in $\ker \varphi$. Let N be the smallest normal subgroup of F containing R. Then since $\ker \varphi$ is a normal subgroup which contains R, $N \subset \ker \varphi$. The mapping property of quo-

tients gives us a homomorphism $\overline{\varphi}: F/N \longrightarrow D_n$. If we show that $\overline{\varphi}$ is bijective, the proposition will be proved.

Note that since φ is surjective, $\overline{\varphi}$ is too. Also, in F/N the relations $\overline{x}^n = 1$, $\overline{y}^2 = 1$, and $\overline{x}\overline{y}\overline{x}\overline{y} = 1$ hold. Using them, we can put any word in $\overline{x}, \overline{y}$ into the form $\overline{x}^i\overline{y}^j$, with $0 \leq i \leq n - 1$ and $0 \leq j \leq 1$. This shows that F/N has at most $2n$ elements. Since $|D_n| = 2n$, it follows that $\overline{\varphi}$ is bijective, as required. □

We will use the notation

$$(8.5) \qquad \langle x_1, \ldots, x_m; r_1, \ldots, r_k \rangle$$

to denote the group generated by elements $x_1, \ldots, x_m$, with defining relations $r_1, \ldots, r_k$. Thus

$$(8.6) \qquad D_n = \langle x, y; x^n, y^2, xyxy \rangle.$$

As a new example, let us consider the group generated by x, y, with the single relation $xyx^{-1}y^{-1} = 1$. If x, y are elements of a group, then

$$(8.7) \qquad xyx^{-1}y^{-1}$$

is called their *commutator*. This commutator is important because it is equal to 1 if and only if x and y commute. This is seen by multiplying both sides of the equation $xyx^{-1}y^{-1} = 1$ on the right by yx. So if we impose the relation $xyx^{-1}y^{-1} = 1$ on the free group, we will obtain a group in which x and y commute. Thus if N is the smallest normal subgroup containing the commutator $xyx^{-1}y^{-1}$ and if $G = F/N$, then the residues of x and y are commuting elements of G. This forces any two elements of G to commute.

(8.8) **Proposition.** Let F be the free group on x, y and let N be the smallest normal subgroup generated by the commutator $xyx^{-1}y^{-1}$. The quotient group $G = F/N$ is abelian.

Proof. Let us denote the residues of the generators x, y in G by the same letters. Since the commutator is in N, the elements x,y commute in G. Then x commutes with y^{-1} too. For xy^{-1} and $y^{-1}x$ both become equal to x when multiplied on the left by y. So by the Cancellation Law, they are equal. Also, x obviously commutes with x and with x^{-1}. So x commutes with any word in $S' = \{x, x^{-1}, y, y^{-1}\}$. So does y. It follows by induction that any two words in S' commute. Since x,y generate the group, G is commutative. □

Note this consequence: The commutator $uvu^{-1}v^{-1}$ of any two words in S' is in the normal subgroup generated by the single commutator $xyx^{-1}y^{-1}$, because, since u, v commute in G, the commutator represents the identity element in G.

The group G constructed above is called the *free abelian group* on the set $\{x, y\}$, because the elements x,y satisfy no relations except those implied by the group axioms and the commutative law.

In the examples we have seen, knowledge of the relations allows us to compute

easily in the group. This is somewhat misleading, because computation with a given set of relations is often not easy at all. For example, suppose that we change the defining relations (8.6) for the dihedral group slightly, substituting y^3 for y^2:

(8.9) $$G = \langle x, y; x^n, y^3, xyxy \rangle.$$

This group is much more complicated. When $n > 5$, it is an infinite group.

Things become very difficult when the relations are complicated enough. Suppose that we are given a set R of words, and let N be the smallest normal subgroup containing R. Let w, w' be any other words. Then we can pose the problem of deciding whether or not w and w' represent the same element of F/N. This is called the *word problem for groups,* and it is known that there is no general procedure for deciding it in a predictable length of time. Nevertheless, generators and relations allow efficient computation in many cases, and so they are a useful tool. We will discuss an important method for computation, the Todd–Coxeter Algorithm, in the next section.

Recapitulating, when we speak of a group defined by generators S and relations R, we mean the quotient group F/N, where F is the free group on S and N is the smallest normal subgroup of F containing R. Note that *any* set R of relations will define a group, because F/N is always defined. The larger R is, the larger N becomes and the more collapsing takes place in the homomorphism $\pi: F \longrightarrow F/N$. If R gets "too big," the worst that can happen is that $N = F$, hence that F/N is the trivial group. Thus there is no such thing as a contradictory set of relations. The only problems which may arise occur when F/N becomes too small, which happens when the relations cause more collapsing than was expected.

9. THE TODD-COXETER ALGORITHM

Let H be a subgroup of a finite group G. The Todd–Coxeter Algorithm which is described in this section is an amazing direct method of counting the cosets of H in G and of determining the operation of G on the set of cosets. Since we know that any operation on an orbit looks like an operation on cosets [Chapter 5 (6.3)], the algorithm is really a method of describing any group operation.

In order to compute explicitly, both the group G and the subgroup H must be given to us in an explicit way. So we consider a group

(9.1) $$G = \langle x_1, \ldots, x_m; r_1, \ldots, r_k \rangle$$

presented by generators $x_1, \ldots, x_m$ and explicitly given relations $r_1, \ldots, r_k$, as in the previous section. Thus G is realized as the quotient group F/N, where F is the free group on the set $\{x_1, \ldots, x_m\}$ and N is the smallest normal subgroup containing $\{r_1, \ldots, r_k\}$. We also assume that the subgroup H of G is given to us explicitly by a set of words

(9.2) $$\{h_1, \ldots, h_s\}$$

in the free group F, whose images in G generate H.

Let us work out a specific example to begin with. We take for G the group generated by three elements x, y, z, with relations x^3, y^2, z^2, xyz, and for H the cyclic subgroup generated by z:

(9.3) $$G = \langle x, y, z; x^3, y^2, z^2, xyz \rangle, \quad H = \{z\}.$$

Since we will be determining the operation on cosets, which is a permutation representation [Chapter 5 (8.1)], we must decide how to write permutations. We will use the cycle notation of Section 6. This forces us to work with *right cosets Hg* rather than with left cosets, because we want G to operate on the right. Let us denote the set of right cosets of H in G by $\mathcal{C}$. We must also decide how to describe the operation of our group explicitly, and the easiest way is to go back to the free group again, that is, to describe the permutations associated to the given generators x, y, z.

The operations of the generators on the set of cosets will satisfy these rules:

(9.4) **Rules.**

1. The operation of each generator (x,y,z in our example) is a permutation.
2. The relations (x^3, y^2, z^2, xyz in our example) operate trivially.
3. The generators of H (z in our example) fix the coset $H1$.
4. The operation on cosets is transitive.

The first rule is a general property of group operations. It follows from the fact that group elements are invertible. We list it instead of mentioning inverses of the generators explicitly. The second rule holds because the relations represent 1 in G, and it is the group G which operates. Rules 3 and 4 are special properties of the operation on cosets.

We now determine the coset representation by applying only these rules. Let us use indices **1**, **2**, **3**,... to denote the cosets, with **1** standing for the coset $H1$. Since we don't know how many cosets there are, we don't know how many indices we need. We will add new ones as necessary.

First, Rule 3 tells us that z sends **1** to itself: $\mathbf{1}z = \mathbf{1}$. This exhausts the information in Rule 3, so Rules 1 and 2 take over. Rule 4 will appear only implicitly.

We don't know what x does to the index **1**. Let's guess that $\mathbf{1}x \neq \mathbf{1}$ and assign a new index, say $\mathbf{1}x = \mathbf{2}$. Continuing with the generator x, we don't know $\mathbf{2}x$, so we assign a third index: $\mathbf{1}x^2 = \mathbf{2}x = \mathbf{3}$. Rule 2 now comes into play. It tells us that x^3 fixes every index. Therefore $\mathbf{1}x^3 = \mathbf{3}x = \mathbf{1}$. It is customary to sum up this information in a table

	x		x		x	
1		**2**		**3**		**1**

which exhibits the operation of x on the three indices. The relation xxx appears on the top, and Rule 2 is reflected in the fact that the same index **1** appears at both ends.

At this point, we have determined the operation of x on the three indices $\mathbf{1}, \mathbf{2}, \mathbf{3}$, except for one thing: We don't yet know that these indices represent distinct cosets.

We now ask for the operation for y on the index $\mathbf{1}$. Again, we don't know it, so we assign a new index, say $\mathbf{1}y = \mathbf{4}$. Rule 2 applies again. Since y^2 operates trivially, we know that $\mathbf{1}y^2 = \mathbf{4}y = \mathbf{1}$:

$$\begin{array}{ccc} & y & y \\ \hline \mathbf{1} & \mathbf{4} & \mathbf{1}. \end{array}$$

The remaining relation is xyz. We know that $\mathbf{1}x = \mathbf{2}$, but we don't yet know $\mathbf{2}y$. So we set $\mathbf{1}xy = \mathbf{2}y = \mathbf{5}$. Rule 2 then tells us that $\mathbf{1}xyz = \mathbf{5}z = \mathbf{1}$:

$$\begin{array}{cccc} & x & y & z \\ \hline \mathbf{1} & \mathbf{2} & \mathbf{5} & \mathbf{1}. \end{array}$$

We now apply Rule 1: The operation of each group element is a permutation of the indices. We have determined that $\mathbf{1}z = \mathbf{1}$ and also that $\mathbf{5}z = \mathbf{1}$. It follows that $\mathbf{5} = \mathbf{1}$. We eliminate the index $\mathbf{5}$, replacing it by $\mathbf{1}$. This in turn tells us that $\mathbf{2}y = \mathbf{1}$. On the other hand, we have already determined that $\mathbf{4}y = \mathbf{1}$. So $\mathbf{4} = \mathbf{2}$ by Rule 1, and we eliminate $\mathbf{4}$.

The entries in the table below have now been determined:

	x	x	x	y	y	z	z	x	y	z
1	**2**	**3**	**1**	**2**	**1**	**1**	**1**	**2**	**1**	**1**
2	**3**	**1**	**2**	**1**	**2**		**2**	**3**		**2**
3	**1**	**2**	**3**		**3**		**3**	**1**	**2**	**3**.

The bottom right corner shows that $\mathbf{2}z = \mathbf{3}$. This determines the rest of the table. There are three indices, and the operation is

$$x = (\mathbf{123}),\ y = (\mathbf{12}),\ z = (\mathbf{23}).$$

Since there are three indices, we conclude that there are three cosets and that the index of H in G is 3. We also conclude that the order of H is 2, and hence that G has order 6. For $z^2 = 1$ is one of our relations; therefore z has order 1 or 2, and since z does not operate trivially on the indices, $z \neq 1$. The three permutations listed above generate the symmetric group, so the permutation representation is an isomorphism from G onto S_3.

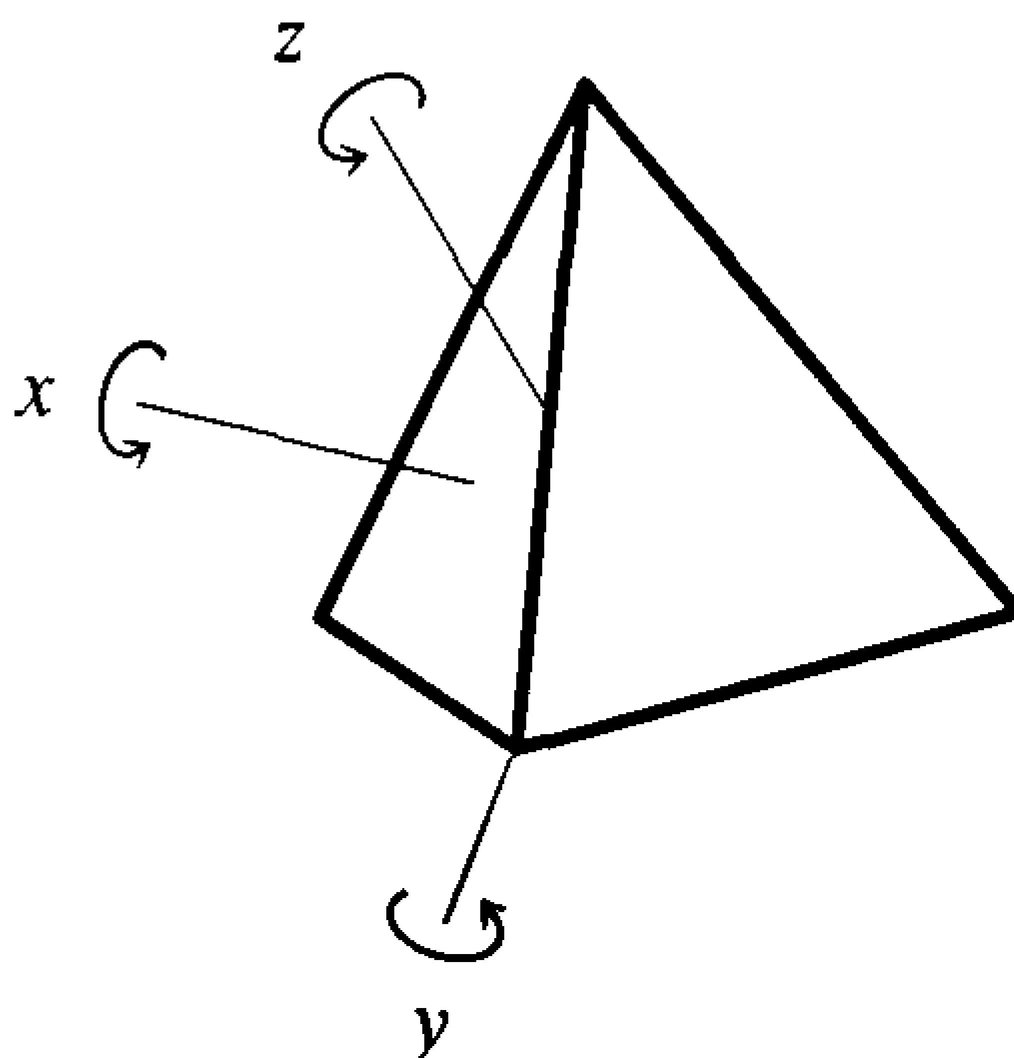
Of course, these conclusions depend on our knowing that the permutation representation we have constructed is the right one. We will show this at the end of the section. Let's compute a few more examples first.

(9.5) **Example.** Consider the tetrahedral group T of the 12 rotational symmetries of a regular tetrahedron (see Section 9 of Chapter 5). If we let y and x denote counterclockwise rotations by $2\pi/3$ about a vertex and the center of a face as shown below,

then $yx = z$ is the rotation by π about an edge. Thus the relations

$$x^3 = 1,\ y^3 = 1,\ yxyx = 1 \tag{9.6}$$

hold in T.



Let us show that (9.6) is a complete set of relations for T. To do so, we consider the group $G = \langle y, x;\ y^3, x^3, yxyx\rangle$ defined by these relations. Since the relations (9.6) hold in T, the mapping property of quotient groups provides a homomorphism $\varphi\colon G \longrightarrow T$. This map is surjective because, as is easily seen, y and x generate T. We need only show that φ is injective. We will do this by showing that the order of the group G is 12.

It is possible to analyze the relations directly, but they aren't particularly easy to work with. We could also compute the order of G by enumerating the cosets of the trivial subgroup $H = \{1\}$. This is not efficient either. It is better to use a nontrivial subgroup H of G, such as the one generated by y. This subgroup has order at most 3 because $y^3 = 1$. If we show that its order is 3 and that its index in G is 4, it will follow that G has order 12, and we will be done.

Here is the resulting table. To fill it in, work from both ends of the relations.

	x	x	x	y	y	y	y	x	y	x
1	2	3	1	1	1	1	1	2	3	1
2	3	1	2	3	4	2	3	1	1	2
3	1	2	3	4	2	3	4	4	2	3
4	4	4	4	2	3	4	2	3	4	4

Thus the permutation representation is

$$x = (\mathbf{1\,2\,3}),\ y = (\mathbf{2\,3\,4}). \tag{9.7}$$

Since there are four indices, the index of H is 4. Also, notice that y does have order precisely 3. For since $y^3 = 1$, the order is at most 3, and since the permutation $(\mathbf{2\,3\,4})$ associated to y has order 3, it is at least 3. So the order of the group is 12, as predicted. Incidentally, we can derive the fact that T is isomorphic to the alternating group A_4 by verifying that the permutations (9.7) generate that group. □

(9.8) **Example.** We modify the relations (9.6) slightly. Let G be generated by x, y, with relations

$$x^3 = 1,\ y^3 = 1,\ yxy^2x = 1,$$

and let H be the subgroup generated by y. Here is a start for a table. Since $y^3 = 1$, we have shortened the last relation, substituting y^{-1} for y^2. Clearly, y^{-1} acts as the inverse of the permutation associated to y. The entries in the bottom row have been determined by working from the right side.

	x	x	x	y	y	y	y	x	y^{-1}	x
1	2	3	**1**	1	1	**1**	**1**	2	3	**1**
2			**2**			**2**	**3**	**1**	**1**	**2**

We rewrite the relation $\mathbf{2}y^{-1} = \mathbf{3}$ as $\mathbf{3}y = \mathbf{2}$. Since $\mathbf{2}y = \mathbf{3}$ as well, it follows that $\mathbf{3}y^2 = \mathbf{3}$ and that $\mathbf{3}y^3 = \mathbf{2}$. But $y^3 = 1$, so $\mathbf{3} = \mathbf{2}$, which in turn implies $\mathbf{1} = \mathbf{2} = \mathbf{3}$. Since the generators x, y fix $\mathbf{1}$, there is one coset, and $H = G$. Therefore x is a power of y. The third relation shows that $x^2 = 1$. Combining this fact with the first relation, we find $x = 1$. Thus G is a cyclic group of order 3. This example illustrates how relations may collapse the group. □

In our examples, we have taken for H the subgroup generated by one of the chosen generators of G, but we could also make the computation with a subgroup H generated by an arbitrary set of words. They must be entered into the computation using Rule 3.

This method can also be used when G is infinite, provided that the index $[G:H]$ is finite. The procedure can not be expected to terminate if there are infinitely many cosets.

We now address the question of why the procedure we have described does give the operation on cosets. A formal proof of this fact is not possible without first defining the algorithm formally, and we have not done this. So we will discuss the question informally. We describe the procedure this way: At a given stage of the computation, we will have some set $\mathbf{I}$ of indices, and the operation of some generators of the group on some indices will have been determined. Let us call this a *partial operation* on $\mathbf{I}$. A partial operation need not be consistent with Rules 1, 2, and 3, but it should be transitive; that is, all indices should be in the "partial orbit" of $\mathbf{1}$. This is where Rule 4 comes in. It tells us not to introduce any indices we don't need.

The starting position is $\mathbf{I} = \{\mathbf{1}\}$, with no operations assigned. At any stage there are two possible steps:

(9.9)

(i) We may equate two indices $\mathbf{i}, \mathbf{j} \in \mathbf{I}$ as a consequence of one of the first three rules, or

(ii) we may choose a generator x and an index $\mathbf{i}$ such that $\mathbf{i}x$ has not yet been determined and define $\mathbf{i}x = \mathbf{j}$, where $\mathbf{j}$ is a new index.

We stop the process when an operation has been determined which is consistent with the rules, that is, when we have a complete, consistent table and the rules hold.

There are two questions to ask: First, will this procedure terminate? Second, if it terminates, is the operation the right one? The answer to both questions is yes. It can be shown that the process always terminates, provided that the group is finite and that preference is given to Step (i). We will not prove this. The more important fact for applications is that if the process terminates, the resulting permutation representation is the right one.

(9.10) **Theorem.** Suppose that a finite number of repetitions of Steps (i) and (ii) yields a consistent table. Then the table defines a permutation representation which is isomorphic, by suitable numbering, to the representation on cosets.

Sketch of proof. Let $\mathbf{I}^*$ denote the final set of indices, with its operation. We will prove the proposition by defining a bijective map $\varphi^*\colon \mathbf{I}^* \longrightarrow \mathscr{C}$ from this set to the set of cosets which is compatible with the two operations. We define φ^* inductively, by defining at each stage a map $\varphi\colon \mathbf{I} \longrightarrow \mathscr{C}$ from the set of indices determined at that stage to $\mathscr{C}$, such that φ is compatible with the partial operation on $\mathbf{I}$. To start, $\{\mathbf{1}\} \longrightarrow \mathscr{C}$ sends $\mathbf{1} \rightsquigarrow H1$. Now suppose that $\varphi\colon \mathbf{I} \longrightarrow \mathscr{C}$ has been defined, and let $\mathbf{I}'$ be the result of applying one of Steps (9.9) to $\mathbf{I}$. In case of Step (ii), there is no difficulty in extending φ to a map $\varphi'\colon \mathbf{I} \longrightarrow \mathscr{C}$. We simply define $\varphi'(\mathbf{k}) = \varphi(\mathbf{k})$ if $\mathbf{k} \neq \mathbf{j}$, and $\varphi'(\mathbf{j}) = \varphi(\mathbf{i})x$. Next, suppose that we use Step (ii) to equate two indices, say $\mathbf{i}, \mathbf{j}$, so that $\mathbf{I}$ is collapsed to form the new index set $\mathbf{I}'$. Then the next lemma allows us to define the map $\varphi'\colon \mathbf{I}' \longrightarrow \mathscr{C}$:

(9.11) **Lemma.** Suppose that a map $\varphi\colon \mathbf{I} \longrightarrow \mathscr{C}$ is given, compatible with a partial operation on $\mathbf{I}$. Let $\mathbf{i}, \mathbf{j} \in \mathbf{I}$, and suppose that one of the Rules 1, 2, or 3 forces $\mathbf{i} = \mathbf{j}$. Then $\varphi(\mathbf{i}) = \varphi(\mathbf{j})$.

Proof. This is true because, as we have already remarked, the operation on cosets does satisfy all of the Rules (9.4). So if the rules force $\mathbf{i} = \mathbf{j}$, they also force $\varphi(\mathbf{i}) = \varphi(\mathbf{j})$. □

It remains to prove that the map $\varphi^*\colon \mathbf{I}^* \longrightarrow \mathscr{C}$ is bijective. To do this, we construct the inverse map $\psi^*\colon \mathscr{C} \longrightarrow \mathbf{I}^*$, using the following lemma:

(9.12) **Lemma.** Let S be a set on which G operates, and let $s \in S$ be an element stabilized by H. There is a unique map $\psi\colon \mathscr{C} \longrightarrow S$ which is compatible with the operations on the two sets and which sends $H1 \rightsquigarrow s$.

Proof. This proof repeats that of (6.4) in Chapter 5, except that we have changed to right operations. Since g sends $H \rightsquigarrow Hg$ and since we want $\psi(Hg) = \psi(H)g$, we must try to set $\psi(Hg) = sg$. This proves uniqueness of the map ψ. To prove existence, we first check that the rule $\psi(Hg) = sg$ is well-defined: If $Ha = Hb$, then $ba^{-1} \in H$. By hypothesis, ba^{-1} stabilizes s, so $sa = sb$. Finally, ψ is compatible with the operations of G because $\psi(Hga) = sga = (sg)a = \psi(Hg)a$. □

Now, to prove the bijectivity of ψ^*, we use the lemma to construct a map $\psi^*: \mathscr{C} \longrightarrow \mathbf{I}^*$. Consider the composed map $\varphi^*\psi^*: \mathscr{C} \longrightarrow \mathscr{C}$. It sends $H1 \rightsquigarrow H1$. We apply the lemma again, substituting $\mathscr{C}$ for S. The uniqueness assertion of the lemma tells us that $\varphi^*\psi^*$ is the identity map. On the other hand, since the operation on $\mathbf{I}^*$ is transitive and since ψ^* is compatible with the operations, ψ^* must be surjective. It follows that φ^* and ψ^* are bijective. □

The axiomatic method has many advantages over honest work.

Bertrand Russell

EXERCISES

1. *The Operations of a Group on Itself*

1. Does the rule $g, x \rightsquigarrow xg^{-1}$ define an operation of G on itself?

2. Let H be a subgroup of a group G. Then H operates on G by left multiplication. Describe the orbits for this operation.

3. Prove the formula $|G| = |Z| + \sum|C|$, where the sum is over the conjugacy classes containing more than one element and where Z is the center of G.

4. Prove the Fixed Point Theorem (1.12).

5. Determine the conjugacy classes in the group M of motions of the plane.

6. Rule out as many of the following as possible as Class Equations for a group of order 10: $1+1+1+2+5$, $1+2+2+5$, $1+2+3+4$, $1+1+2+2+2+2$.

7. Let $F = \mathbb{F}_5$. Determine the order of the conjugacy class of $\begin{bmatrix} 1 & \\ & 2 \end{bmatrix}$ in $GL_2(\mathbb{F}_5)$.

8. Determine the Class Equation for each of the following groups.
(a) the quaternion group, **(b)** the Klein four group, **(c)** the dihedral group D_5,
(d) D_6, **(e)** D_n, **(f)** the group of upper triangular matrices in $GL_2(\mathbb{F}_3)$,
(g) $SL_2(\mathbb{F}_3)$.

9. Let G be a group of order n, and let F be any field. Prove that G is isomorphic to a subgroup of $GL_n(F)$.

10. Determine the centralizer in $GL_3(\mathbb{R})$ of each matrix.

(a) $\begin{bmatrix} 1 & & \\ & 2 & \\ & & 3 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & & \\ & 1 & \\ & & 2 \end{bmatrix}$ **(c)** $\begin{bmatrix} 1 & 1 & \\ & 1 & \\ & & 1 \end{bmatrix}$ **(d)** $\begin{bmatrix} 1 & 1 & \\ & 1 & 1 \\ & & 1 \end{bmatrix}$

(e) $\begin{bmatrix} 1 & & \\ & & 1 \\ & 1 & \end{bmatrix}$ **(f)** $\begin{bmatrix} & 1 & \\ & & 1 \\ 1 & & \end{bmatrix}$

***11.** Determine all finite groups which contain at most three conjugacy classes.

12. Let N be a normal subgroup of a group G. Suppose that $|N| = 5$ and that $|G|$ is odd. Prove that N is contained in the center of G.

*13. (a) Determine the possible Class Equations for groups of order 8.
(b) Classify groups of order 8.

14. Let Z be the center of a group G. Prove that if G/Z is a cyclic group, then G is abelian and hence $G = Z$.

*15. Let G be a group of order 35.
(a) Suppose that G operates nontrivially on a set of five elements. Prove that G has a normal subgroup of order 7.
(b) Prove that every group of order 35 is cyclic.

2. The Class Equation of the Icosahedral Group

1. Identify the intersection $I \cap O$ when the dodecahedron and cube are as in Figure (2.7).
2. Two tetrahedra can be inscribed into a cube C, each one using half the vertices. Relate this to the inclusion $A_4 \subset S_4$.
3. Does I contain a subgroup T? D_6? D_3?
4. Prove that the icosahedral group has no subgroup of order 30.
5. Prove or disprove: A_5 is the only proper normal subgroup of S_5.
6. Prove that no group of order p^e, where p is prime and $e > 1$, is simple.
7. Prove or disprove: An abelian group is simple if and only if it has prime order.
8. (a) Determine the Class Equation for the group T of rotations of a tetrahedron.
(b) What is the center of T?
(c) Prove that T has exactly one subgroup of order 4.
(d) Prove that T has no subgroup of order 6.
9. (a) Determine the Class Equation for the octahedral group O.
(b) There are exactly two proper normal subgroups of O. Find them, show that they are normal, and show that there are no others.
10. Prove that the tetrahedral group T is isomorphic to the alternating group A_4, and that the octahedral group O is isomorphic to the symmetric group S_4. Begin by finding sets of four elements on which these groups operate.
11. Prove or disprove: The icosahedral group is not a subgroup of the group of real upper triangular 2×2 matrices.

*12. Prove or disprove: A nonabelian simple group can not operate nontrivially on a set containing fewer than five elements.

3. Operations on Subsets

1. Let S be the set of subsets of order 2 of the dihedral group D_3. Determine the orbits for the action of D_3 on S by conjugation.
2. Determine the orbits for left multiplication and for conjugation on the set of subsets of order 3 of D_3.
3. List all subgroups of the dihedral group D_4, and divide them into conjugacy classes.
4. Let H be a subgroup of a group G. Prove that the orbit of the left coset gH for the operation of conjugation contains the right coset Hg.
5. Let U be a subset of a finite group G, and suppose that $|U|$ and $|G|$ have no common factor. Is the stabilizer of $|U|$ trivial for the operation of conjugation?
6. Consider the operation of left multiplication by G on the set of its subsets. Let U be a

subset whose orbit $\{gU\}$ partitions G. Let H be the unique subset in this orbit which contains 1. Prove that H is a subgroup of G and that the sets gU are its left cosets.

7. Let H be a subgroup of a group G. Prove or disprove: The normalizer $N(H)$ is a normal subgroup of the group G.

8. Let $H \subset K \subset G$ be groups. Prove that H is normal in K if and only if $K \subset N(H)$.

9. Prove that the subgroup B of upper triangular matrices in $GL_n(\mathbb{R})$ is conjugate to the group L of lower triangular matrices.

10. Let B be the subgroup of $G = GL_n(\mathbb{C})$ of upper triangular matrices, and let $U \subset B$ be the set of upper triangular matrices with diagonal entries 1. Prove that $B = N(U)$ and that $B = N(B)$.

***11.** Let S_n denote the subgroup of $GL_n(\mathbb{R})$ of permutation matrices. Determine the normalizer of S_n in $GL_n(\mathbb{R})$.

12. Let S be a finite set on which a group G operates transitively, and let U be a subset of S. Prove that the subsets gU cover S evenly, that is, that every element of S is in the same number of sets gU.

13. **(a)** Let H be a normal subgroup of G of order 2. Prove that H is in the center of G.
(b) Let H be a normal subgroup of prime order p in a finite group G. Suppose that p is the smallest prime dividing $|G|$. Prove that H is in the center $Z(G)$.

***14.** Let H be a proper subgroup of a finite group G. Prove that the union of the conjugates of H is not the whole group G.

15. Let K be a normal subgroup of order 2 of a group G, and let $\overline{G} = G/K$. Let $\overline{C}$ be a conjugacy class in $\overline{G}$. Let S be the inverse image of $\overline{C}$ in G. Prove that one of the following two cases occurs.
(a) $S = C$ is a single conjugacy class and $|C| = 2|\overline{C}|$.
(b) $S = C_1 \cup C_2$ is made up of two conjugacy classes and $|C_1| = |C_2| = |\overline{C}|$.

16. Calculate the **double cosets** HgH of the subgroup $H = \{1, y\}$ in the dihedral group D_n. Show that each double coset has either two or four elements.

17. Let H, K be subgroups of G, and let H' be a conjugate subgroup of H. Relate the double cosets $H'gK$ and HgK.

18. What can you say about the order of a double coset HgK?

4. The Sylow Theorems

1. How many elements of order 5 are contained in a group of order 20?

2. Prove that no group of order pq, where p and q are prime, is simple.

3. Prove that no group of order p^2q, where p and q are prime, is simple.

4. Prove that the set of matrices $\begin{bmatrix} 1 & a \\ & c \end{bmatrix}$ where $a, c \in \mathbb{F}_7$ and $c = 1, 2, 4$ forms a group of the type presented in (4.9b) (and that therefore such a group exists).

5. Find Sylow 2-subgroups in the following cases:
(a) D_{10} **(b)** T **(c)** O **(d)** I.

6. Find a Sylow p-subgroup of $GL_2(\mathbb{F}_p)$.

***7.** **(a)** Let H be a subgroup of G of prime index p. What are the possible numbers of conjugate subgroups of H?
(b) Suppose that p is the smallest prime integer which divides $|G|$. Prove that H is a normal subgroup.

*8. Let H be a Sylow p-su group of G, and let $K = N(H)$. Prove or disprove: $K = N(K)$.

9. Let G be a group of order $p^e m$. Prove that G contains a subgroup of order p^r for every integer $r \leq e$.

10. Let $n = pm$ be an integer which is divisible exactly once by p, and let G be a group of order n. Let H be a Sylow p-subgroup of G, and let S be the set of all Sylow p-subgroups. How does S decompose into H-orbits?

*11. (a) Compute the order of $GL_n(\mathbb{F}_p)$.
(b) Find a Sylow p-subgroup of $GL_n(\mathbb{F}_p)$.
(c) Compute the number of Sylow p-subgroups.
(d) Use the Second Sylow Theorem to give another proof of the First Sylow Theorem.

*12. Prove that no group of order 224 is simple.

13. Prove that if G has order $n = p^e a$ where $1 \leq a < p$ and $e \geq 1$, then G has a proper normal subgroup.

14. Prove that the only simple groups of order < 60 are groups of prime order.

15. Classify groups of order 33.

16. Classify groups of order 18.

17. Prove that there are at most five isomorphism classes of groups of order 20.

*18. Let G be a simple group of order 60.
(a) Prove that G contains six Sylow 5-subgroups, ten Sylow 3-subgroups, and five Sylow 2-subgroups.
(b) Prove that G is isomorphic to the alternating group A_5.

5. The Groups of Order 12

1. Determine the Class Equations of the groups of order 12.

2. Prove that a group of order $n = 2p$, where p is prime, is either cyclic or dihedral.

*3. Let G be a group of order 30.
(a) Prove that either the Sylow 5-subgroup K or the Sylow 3-subgroup H is normal.
(b) Prove that HK is a cyclic subgroup of G.
(c) Classify groups of order 30.

4. Let G be a group of order 55.
(a) Prove that G is generated by two elements x,y, with the relations $x^{11} = 1$, $y^5 = 1$, $yxy^{-1} = x^r$, for some r, $1 \leq r < 11$.
(b) Prove that the following values of r are not possible: 2, 6, 7, 8, 10.
(c) Prove that the remaining values are possible, and that there are two isomorphism classes of groups of order 55.

6. Computation in the Symmetric Group

1. Verify the products (6.9).

2. Prove explicitly that the permutation $(\mathbf{1\,2\,3})(\mathbf{4\,5})$ is conjugate to $(\mathbf{2\,4\,1})(\mathbf{3\,5})$.

3. Let p, q be permutations. Prove that the products pq and qp have cycles of equal sizes.

4. (a) Does the symmetric group S_7 contain an element of order 5? of order 10? of order 15?
(b) What is the largest possible order of an element of S_7?

5. Show how to determine whether a permutation is odd or even when it is written as a product of cycles.
6. Prove or disprove: The order of a permutation is the least common multiple of the orders of the cycles which make it up.
7. Is the cyclic subgroup H of S_n generated by the cycle $(\mathbf{12345})$ a normal subgroup?
*8. Compute the number of permutations in S_n which do not leave any index fixed.
9. Determine the cycle decomposition of the permutation $\mathbf{i} \rightsquigarrow \mathbf{n-i}$.
10. (a) Prove that every permutation p is a product of transpositions.
    (b) How many transpositions are required to write the cycle $(\mathbf{123\cdots n})$?
    (c) Suppose that a permutation is written in two ways as a product of transpositions, say $p = \tau_1\tau_2\cdots\tau_m$ and $p = \tau_1'\tau_2'\cdots\tau_n'$. Prove that m and n are both odd or else they are both even.
11. What is the centralizer of the element $(\mathbf{12})$ of S_4?
12. Find all subgroups of order 4 of the symmetric group S_4. Which are normal?
13. Determine the Class Equation of A_4.
14. (a) Determine the number of conjugacy classes and the Class Equation for S_5.
    (b) List the conjugacy classes in A_5, and reconcile this list with the list of conjugacy classes in the icosahedral group [see (2.2)].
15. Prove that the transpositions $(\mathbf{12}), (\mathbf{23}), \ldots, (\mathbf{n-1}, \mathbf{n})$ generate the symmetric group S_n.
16. Prove that the symmetric group S_n is generated by the cycles $(\mathbf{12\cdots n})$ and $(\mathbf{12})$.
17. (a) Show that the product of two transpositions $(\mathbf{ij})(\mathbf{kl})$ can always be written as a product of 3-cycles. Treat the case that some indices are equal too.
    (b) Prove that the alternating group A_n is generated by 3-cycles, if $n \geq 3$.
18. Prove that if a proper normal subgroup of S_n contains a 3-cycle, it is A_n.
*19. Prove that A_n is simple for all $n \geq 5$.
*20. Prove that A_n is the only subgroup of S_n of index 2.
21. Explain the miraculous coincidence at the end of the section in terms of the opposite group (Chapter 2, Section 1, exercise 12).

7. *The Free Group*

1. Prove or disprove: The free group on two generators is isomorphic to the product of two infinite cyclic groups.
2. (a) Let F be the free group on x,y. Prove that the two elements $u = x^2$ and $v = y^3$ generate a subgroup of F which is isomorphic to the free group on u, v.
   (b) Prove that the three elements $u = x^2$, $v = y^2$, and $z = xy$ generate a subgroup isomorphic to the free group on u, v, z.
3. We may define a *closed word* in S' to be the oriented loop obtained by joining the ends of a word. Thus

$$\begin{array}{ccc} & ca^{-1} & \\ b & & b^{-1} \\ a & & b, \\ a & & c \\ & bbd & \end{array}$$

represents a closed word, if we read it clockwise. Establish a bijective correspondence between reduced closed words and conjugacy classes in the free group.

4. Let p be a prime integer. Let N be the number of words of length p in a finite set S. Show that N is divisible by p.

8. Generators and Relations

1. Prove that two elements a, b of a group generate the same subgroup as bab^2, bab^3.
2. Prove that the smallest normal subgroup of a group G containing a subset S is generated as a subgroup by the set $\{gsg^{-1} \mid g \in G, s \in S\}$.
3. Prove or disprove: y^2x^2 is in the normal subgroup generated by xy and its conjugates.
4. Prove that the group generated by x, y, z with the single relation $yxyz^{-2} = 1$ is actually a free group.
5. Let S be a set of elements of a group G, and let $\{r_i\}$ be some relations which hold among the elements S in G. Let F be the free group on S. Prove that the map $F \longrightarrow G$ (8.1) factors through F/N, where N is the normal subgroup generated by $\{r_i\}$.
6. Let G be a group with a normal subgroup N. Assume that G and G/N are both cyclic groups. Prove that G can be generated by two elements.
7. A subgroup H of a group G is called *characteristic* if it is carried to itself by all automorphisms of G.
   (a) Prove that every characteristic subgroup is normal.
   (b) Prove that the center Z of a group G is a characteristic subgroup.
   (c) Prove that the subgroup H generated by all elements of G of order n is characteristic.
8. Determine the normal subgroups and the characteristic subgroups of the quaternion group.
9. The *commutator subgroup* C of a group G is the smallest subgroup containing all commutators.
   (a) Prove that the commutator subgroup is a characteristic subgroup.
   (b) Prove that G/C is an abelian group.
10. Determine the commutator subgroup of the group M of motions of the plane.
11. Prove by explicit computation that the commutator $x(yz)x^{-1}(yz)^{-1}$ is in the normal subgroup generated by the two commutators $xyx^{-1}y^{-1}$ and $xzx^{-1}z^{-1}$ and their conjugates.
12. Let G denote the free abelian group $\langle x, y; xyx^{-1}y^{-1}\rangle$ defined in (8.8). Prove the universal property of this group: If u, v are elements of an abelian group A, there is a unique homomorphism $\varphi: G \longrightarrow A$ such that $\varphi(x) = u$, $\varphi(y) = v$.
13. Prove that the normal subgroup in the free group $\langle x, y\rangle$ which is generated by the single commutator $xyx^{-1}y^{-1}$ is the commutator subgroup.
14. Let N be a normal subgroup of a group G. Prove that G/N is abelian if and only if N contains the commutator subgroup of G.
15. Let $\varphi: G \longrightarrow G'$ be a surjective group homomorphism. Let S be a subset of G such that $\varphi(S)$ generates G', and let T be a set of generators of $\ker \varphi$. Prove that $S \cup T$ generates G.
16. Prove or disprove: Every finite group G can be presented by a finite set of generators and a finite set of relations.
17. Let G be the group generated by x, y, z, with certain relations $\{r_i\}$. Suppose that one of the relations has the form wx, where w is a word in y, z. Let r_i' be the relation obtained by substituting w^{-1} for x into r_i, and let G' be the group generated by y, z, with relations $\{r_i'\}$. Prove that G and G' are isomorphic.

9. The Todd–Coxeter Algorithm

1. Prove that the elements x, y of (9.5) generate T, and that the permutations (9.7) generate A_4.
2. Use the Todd–Coxeter Algorithm to identify the group generated by two elements x, y, with the following relations.
   (a) $x^2 = y^2 = 1, xyx = yxy$
   (b) $x^2 = y^3 = 1, xyx = yxy$
   (c) $x^3 = y^3 = 1, xyx = yxy$
   (d) $x^4 = y^2 = 1, xyx = yxy$
   (e) $x^4 = y^4 = x^2y^2 = 1$
3. Use the Todd–Coxeter Algorithm to determine the order of the group generated by x, y, with the following relations.
   (a) $x^4 = 1, y^3 = 1, xy = y^2x$ **(b)** $x^7 = 1, y^3 = 1, yx = x^2y$.
4. Identify the group G generated by elements x, y, z, with relations $x^4 = y^4 = z^3 = x^2z^2 = 1$ and $z = xy$.
5. Analyze the group G generated by x, y, with relations $x^4 = 1$, $y^4 = 1$, $x^2 = y^2$, $xy = y^3x$.
*6. Analyze the group generated by elements x, y, with relations $x^{-1}yx = y^{-1}$, $y^{-1}xy = x^{-1}$.
7. Let G be the group generated by elements x, y, with relations $x^4 = 1, y^3 = 1, x^2 = yxy$. Prove that this group is trivial in these two ways.
   (a) using the Todd–Coxeter Algorithm
   (b) working directly with the relations
8. Identify the group G generated by two elements x, y, with relations $x^3 = y^3 = yxyxy = 1$.
9. Let $p \le q \le r$ be integers >1. The *triangle group* G^{pqr} is defined by generators $G^{pqr} = \langle x, y, z; x^p, y^q, z^r, xyz \rangle$. In each case, prove that the triangle group is isomorphic to the group listed.
   (a) the dihedral group D_n, when $p, q, r = 2, 2, n$
   (b) the tetrahedral group, when $p, q, r = 2, 3, 3$
   (c) the octahedral group, when $p, q, r = 2, 3, 4$
   (d) the icosahedral group, when $p, q, r = 2, 3, 5$
10. Let Δ denote an isosceles right triangle, and let a, b, c denote the reflections of the plane about the three sides of Δ. Let $x = ab$, $y = bc$, $z = ca$. Prove that x, y, z generate a triangle group.
11. **(a)** Prove that the group G generated by elements x, y, z with relations $x^2 = y^3 = z^5 = 1, xyz = 1$ has order 60.
    (b) Let H be the subgroup generated by x and zyz^{-1}. Determine the permutation representation of G on G/H, and identify H.
    (c) Prove that G is isomorphic to the alternating group A_5.
    (d) Let K be the subgroup of G generated by x and yxz. Determine the permutation representation of G on G/K, and identify K.

Miscellaneous Problems

1. **(a)** Prove that the subgroup T' of O_3 of all symmetries of a regular tetrahedron, including orientation-reversing symmetries, has order 24.

(b) Is T' isomorphic to the symmetric group S_4?
(c) State and prove analogous results for the group of symmetries of a dodecahedron.

2. (a) Let $U = \{1, x\}$ be a subset of order 2 of a group G. Consider the graph having one vertex for each element of G and an edge joining the vertices g to gx for all $g \in G$. Prove that the vertices connected to the vertex 1 are the elements of the cyclic group generated by x.
(b) Do the analogous thing for the set $U = \{1, x, y\}$.

*3. (a) Suppose that a group G operates transitively on a set S, and that H is the stabilizer of an element $s_0 \in S$. Consider the action of G on $S \times S$ defined by $g(s_1, s_2) = (gs_1, gs_2)$. Establish a bijective correspondence between double cosets of H in G and G-orbits in $S \times S$.
(b) Work out the correspondence explicitly for the case that G is the dihedral group D_5 and S is the set of vertices of a 5-gon.
(c) Work it out for the case that $G = T$ and that S is the set of edges of a tetrahedron.

*4. Assume that $H \subset K \subset G$ are subgroups, that H is normal in K, and that K is normal in G. Prove or disprove: H is normal in G.

*5. Prove the *Bruhat decomposition,* which asserts that $GL_n(\mathbb{R})$ is the union of the double cosets BPB, where B is the group of upper triangular matrices and P is a permutation matrix.

6. (a) Prove that the group generated by x, y with relations x^2, y^2 is an infinite group in two ways:
(i) It is clear that every word can be reduced by using these relations to the form $\cdots xyxy \cdots$. Prove that every element of G is represented by exactly one such word.
(ii) Exhibit G as the group generated by reflections r, r' about lines ℓ, ℓ' whose angle of intersection is not a rational multiple of 2π.
(b) Let N be any proper normal subgroup of G. Prove that G/N is a dihedral group.

7. Let H, N be subgroups of a group G, and assume that N is a normal subgroup.
(a) Determine the kernels of the restrictions of the canonical homomorphism $\pi: G \longrightarrow G/N$ to the subgroups H and HN.
(b) Apply the First Isomorphism Theorem to these restrictions to prove the *Second Isomorphism Theorem:* $H/(H \cap N)$ is isomorphic to $(HN)/N$.

8. Let H, N be normal subgroups of a group G such that $H \supset N$, and let $\overline{H} = H/N$, $\overline{G} = G/N$.
(a) Prove that $\overline{H}$ is a normal subgroup of $\overline{G}$.
(b) Use the composed homomorphism $G \longrightarrow \overline{G} \longrightarrow \overline{G}/\overline{H}$ to prove the *Third Isomorphism Theorem:* G/H is isomorphic to $\overline{G}/\overline{H}$.

Chapter 7

Bilinear Forms

I presume that to the uninitiated
the formulae will appear cold and cheerless.

Benjamin Pierce

1. DEFINITION OF BILINEAR FORM

Our model for bilinear forms is the dot product

$$(1.1) \qquad (X \cdot Y) = X^t Y = x_1 y_1 + \cdots + x_n y_n$$

of vectors in $\mathbb{R}^n$, which was described in Section 5 of Chapter 4. The symbol $(X \cdot Y)$ has various properties, the most important for us being the following:

(1.2) *Bilinearity:*

$$(X_1 + X_2 \cdot Y) = (X_1 \cdot Y) + (X_2 \cdot Y)$$
$$(X \cdot Y_1 + Y_2) = (X \cdot Y_1) + (X \cdot Y_2)$$
$$(cX \cdot Y) = c(X \cdot Y) = (X \cdot cY)$$

Symmetry:

$$(X \cdot Y) = (Y \cdot X)$$

Positivity:

$$(X \cdot X) > 0, \quad \text{if } X \neq 0.$$

Notice that bilinearity says this: If one variable is fixed, the resulting function of the remaining variable is a linear transformation $\mathbb{R}^n \longrightarrow \mathbb{R}$.

We will study dot product and its analogues in this chapter. It is clear how to generalize bilinearity and symmetry to a vector space over any field, while positivity is, a priori, applicable only when the scalar field is $\mathbb{R}$. We will also extend the concept of positivity to complex vector spaces in Section 4.

Let V be a vector space over a field F. A *bilinear form* on V is a function of two variables on V, with values in the field: $V \times V \xrightarrow{f} F$, satisfying the bilinear axioms, which are

$$(1.3)\qquad \begin{aligned} f(v_1 + v_2, w) &= f(v_1, w) + f(v_2, w) \\ f(cv, w) &= cf(v, w) \\ f(v, w_1 + w_2) &= f(v, w_1) + f(v, w_2) \\ f(v, cw) &= cf(v, w) \end{aligned}$$

for all $v, w, v_i, w_i \in V$ and all $c \in F$. Often a notation similar to dot product is used. We will frequently use the notation

$$(1.4)\qquad \langle v, w \rangle$$

to designate the value $f(u, v)$ of the form. So $\langle v, w \rangle$ is a scalar, an element of F.

A form $\langle\, ,\rangle$ is said to be *symmetric* if

$$(1.5)\qquad \langle v, w \rangle = \langle w, v \rangle$$

and *skew-symmetric* if

$$(1.6)\qquad \langle v, w \rangle = -\langle w, v \rangle,$$

for all $v, w \in V$. (This is actually not the right definition of skew-symmetry if the field F is of characteristic 2, that is, if $1 + 1 = 0$ in F. We will correct the definition in Section 8.)

If the form f is either symmetric or skew-symmetric, then linearity in the second variable follows from linearity in the first.

The main examples of bilinear forms are the forms on the space F^n of column vectors, obtained as follows: Let A be an $n \times n$ matrix in F, and define

$$(1.7)\qquad \langle X, Y \rangle = X^t A Y.$$

Note that this product is a 1×1 matrix, that is, a scalar, and that it is bilinear. Ordinary dot product is included as the case $A = I$.

A matrix A is *symmetric* if

$$(1.8)\qquad A^t = A, \quad \text{that is, } a_{ij} = a_{ji} \quad \text{for all } i, j.$$

(1.9) Proposition. The form (1.7) is symmetric if and only if the matrix A is symmetric.

Proof. Assume that A is symmetric. Since Y^tAX is a 1×1 matrix, it is equal to its transpose: $Y^tAX = (Y^tAX)^t = X^tA^tY = X^tAY$. Thus $\langle Y, X \rangle = \langle X, Y \rangle$. The other implication is obtained by setting $X = e_i$ and $Y = e_j$. We find $\langle e_i, e_j \rangle = e_i{}^tAe_j = a_{ij}$, while $\langle e_j, e_i \rangle = a_{ji}$. If the form is symmetric, then $a_{ij} = a_{ji}$, and so A is symmetric. □

Let $\langle\, ,\rangle$ be a bilinear form on a vector space V, and let $\mathbf{B} = (v_1, \ldots, v_n)$ be a basis for V. We can relate the form to a product X^tAY by the *matrix of the form* with

respect to the basis. By definition, this is the matrix $A = (a_{ij})$, where

$$(1.10) \qquad a_{ij} = \langle v_i, v_j \rangle.$$

Note that A is a symmetric matrix if and only if $\langle\, ,\rangle$ is a symmetric form. Also, the symmetry of the bilinear form does not depend on the basis. So if the matrix of the form with respect to some basis is symmetric, its matrix with respect to any other basis will be symmetric too.

The matrix A allows us to compute the value of the form on two vectors $v, w \in V$. Let X, Y be their coordinate vectors, as in Section 4 of Chapter 3, so that $v = \mathbf{B}X$, $w = \mathbf{B}Y$. Then

$$\langle v, w \rangle = \langle \sum_i v_i x_i, \sum_j v_j y_j \rangle.$$

This expands using bilinearity to $\sum_{i,j} x_i y_j \langle v_i, v_j \rangle = \sum_{i,j} x_i a_{ij} y_j = X^t A Y$:

$$(1.11) \qquad \langle v, w \rangle = X^t A Y.$$

Thus, if we identify F^n with V using the basis $\mathbf{B}$ as in Chapter 3 (4.14), the bilinear form $\langle\, ,\rangle$ corresponds to $X^t A Y$.

As in the study of linear operators, a central problem is to describe the effect of a change of basis on such a product. For example, we would like to know what happens to dot product when the basis of $\mathbb{R}^n$ is changed. This will be discussed presently. The effect of a change of basis $\mathbf{B} = \mathbf{B}'P$ [Chapter 3 (4.16)] on the matrix of the form can be determined easily from the rules $X' = PX$, $Y' = PY$: If A' is the matrix of the form with respect to a new basis $\mathbf{B}'$, then by definition of A', $\langle v, w \rangle = X'^t A' Y' = X^t P^t A' P Y$. But we also have $\langle v, w \rangle = X^t A Y$. So

$$(1.12) \qquad P^t A' P = A.$$

Let $Q = (P^{-1})^t$. Since P can be any invertible matrix, Q is also arbitrary.

(1.13) **Corollary.** Let A be the matrix of a bilinear form with respect to a basis. The matrices A' which represent the same form with respect to different bases are the matrices $A' = QAQ^t$, where Q is an arbitrary matrix in $GL_n(F)$. □

Let us now apply formula (1.12) to our original example of dot product on $\mathbb{R}^n$. The matrix of the dot product with respect to the standard basis is the identity matrix: $(X \cdot Y) = X^t I Y$. So formula (1.12) tells us that if we change basis, the matrix of the form changes to

$$(1.14) \qquad A' = (P^{-1})^t I (P^{-1}) = (P^{-1})^t (P^{-1}),$$

where P is the matrix of change of basis as before. If the matrix P happens to be *orthogonal,* meaning that $P^t P = I$, then $A' = I$, and dot product carries over to dot product: $(X \cdot Y) = (PX \cdot PY) = (X' \cdot Y')$, as we saw in Chapter 4 (5.13). But under a general change of basis, the formula for dot product changes to $X'^t A' Y'$, where A' is

as in (1.14). For example, let $n = 2$, and let the basis $\mathbf{B}'$ be

$$v_1' = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad v_2' = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Then

$$(1.15) \qquad P^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad A' = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

The matrix A' represents dot product on $\mathbb{R}^2$, with respect to the basis $\mathbf{B}'$.

We can also turn the computation around. Suppose that we are given a bilinear form $\langle\, ,\rangle$ on a real vector space V. Let us ask whether or not this form becomes dot product when we choose a suitable basis. We start with an arbitrary basis $\mathbf{B}$, so that we have a matrix A to work with. Then the problem is to change this basis in such a way that the new matrix is the identity, if that is possible. By formula (1.12), this amounts to solving the matrix equation $I = (P^{-1})^t A (P^{-1})$, or

$$(1.16) \qquad A = P^t P.$$

(1.17) **Corollary.** The matrices A which represent a form equivalent to dot product are the matrices $A = P^t P$, where P is invertible. □

This corollary gives a theoretical answer to our problem of determining the bilinear forms equivalent to dot product, but it is not very satisfactory because we don't yet have a practical method of deciding which matrices can be written as a product $P^t P$, let alone a practical method of finding P.

We can get some conditions on the matrix A from the properties of dot product listed in (1.2). Bilinearity imposes no condition on A, because the symbol $X^t A Y$ is always bilinear. However, symmetry and positivity restrict the possibilities. The easier property to check is symmetry: In order to represent dot product, the matrix A must be symmetric. Positivity is also a strong restriction. In order to represent dot product, the matrix A must have the property that

$$(1.18) \qquad X^t A X > 0, \quad \text{for all } X \neq 0.$$

A real symmetric matrix having this property is called *positive definite*.

(1.19) **Theorem.** The following properties of a real $n \times n$ matrix A are equivalent:

(i) A represents dot product, with respect to some basis of $\mathbb{R}^n$.
(ii) There is an invertible matrix $P \in GL_n(\mathbb{R})$ such that $A = P^t P$.
(iii) A is symmetric and positive definite.

We have seen that (i) and (ii) are equivalent [Corollary (1.17)] and that (i) implies (iii). So it remains to prove the remaining implication, that (iii) implies (i). It will be more convenient to restate this implication in vector space form.

A symmetric bilinear form $\langle\, ,\rangle$ on a finite-dimensional real vector space V is called *positive definite* if

$$\langle v, v\rangle > 0 \tag{1.20}$$

for every nonzero vector $v \in V$. Thus a real symmetric matrix A is positive definite if and only if the form $\langle X, Y\rangle = X^tAY$ it defines on $\mathbb{R}^n$ is a positive definite form. Also, the form $\langle\, ,\rangle$ is positive definite if and only if its matrix A with respect to any basis is a positive definite matrix. This is clear, because if X is the coordinate vector of a vector v, then $\langle v, v\rangle = X^tAX$ (1.11).

Two vectors v, w are called *orthogonal* with respect to a symmetric form if $\langle v, w\rangle = 0$. Orthogonality of two vectors is often denoted as

$$v \perp w. \tag{1.21}$$

This definition extends the concept of orthogonality which we have already seen when the form is dot product on $\mathbb{R}^n$ [Chapter 4 (5.12)]. A basis $\mathbf{B} = (v_1,\dots,v_n)$ of V is called an *orthonormal basis* with respect to the form if

$$\langle v_i, v_j\rangle = 0 \quad \text{for all } i \neq j, \text{ and } \quad \langle v_i, v_i\rangle = 1 \quad \text{for all } i.$$

It follows directly from the definition that a basis $\mathbf{B}$ is orthonormal if and only if the matrix of the form with respect to $\mathbf{B}$ is the identity matrix.

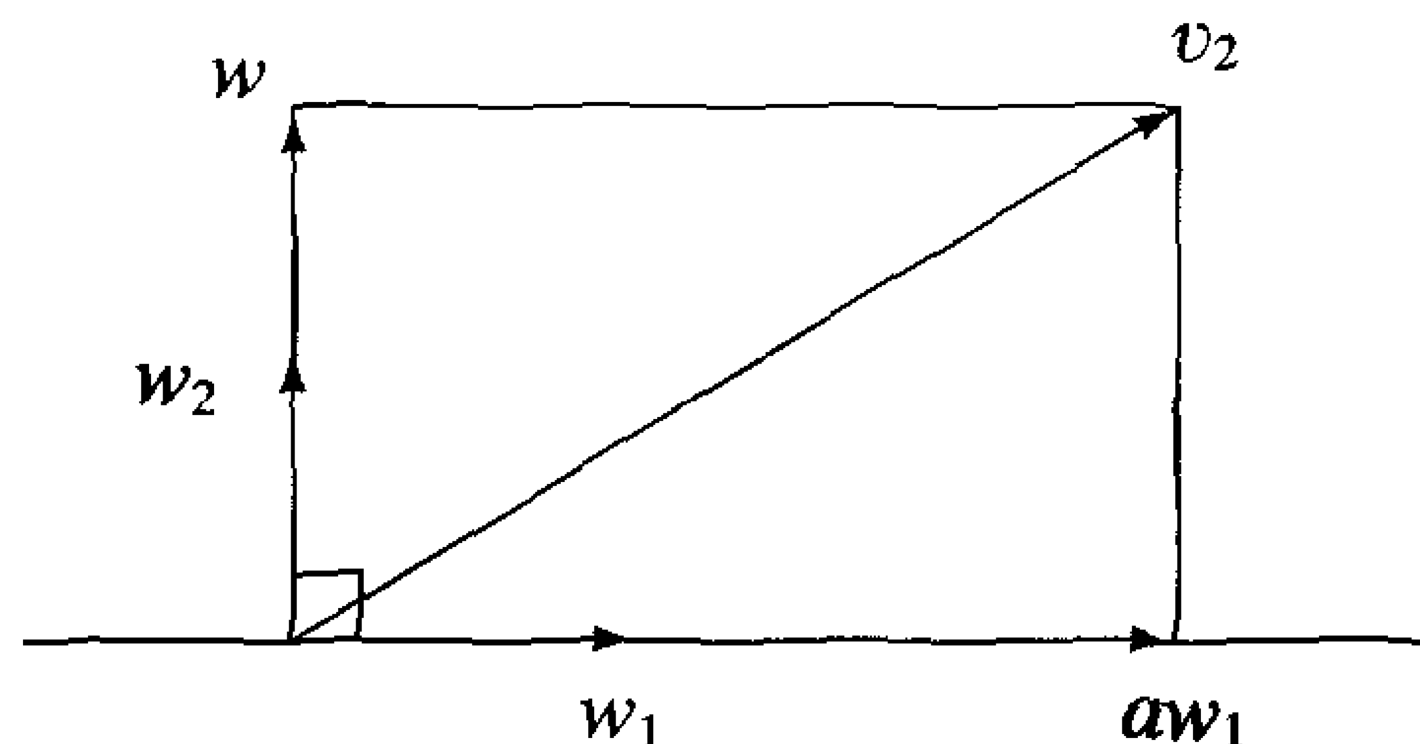
(1.22) **Theorem.** Let $\langle\, ,\rangle$ be a positive definite symmetric form on a finite-dimensional vector real space V. There exists an orthonormal basis for V.

Proof. We will describe a method called the *Gram–Schmidt procedure* for constructing an orthonormal basis, starting with an arbitrary basis $\mathbf{B} = (v_1,\dots,v_n)$. Our first step is to normalize v_1, so that $\langle v_1, v_1\rangle = 1$. To do this we note that

$$\langle cv, cv\rangle = c^2v. \tag{1.23}$$

Since the form is positive definite, $\langle v_1, v_1\rangle > 0$. We set $c = \langle v_1, v_1\rangle^{-\frac{1}{2}}$, and replace v_1 by $w_1 = cv_1$.

Next we look for a linear combination of w_1 and v_2 which is orthogonal to w_1. The required linear combination is $w = v_2 - aw_1$, where $a = \langle v_2, w_1\rangle : \langle w, w_1\rangle = \langle v_2, w_1\rangle - a\langle w_1, w_1\rangle = \langle v_2, w_1\rangle - a = 0$. We normalize this vector w to length 1. obtaining a vector w_2 which we substitute for v_2. The geometric interpretation of this operation is illustrated below for the case that the form is dot product. The vector aw_1 is the orthogonal projection of v_2 onto the subspace (the line) spanned by w_1.



This is the general procedure. Suppose that the $k-1$ vectors $w_1,\dots,w_{k-1}$ are orthonormal and that $(w_1,\dots,w_{k-1},v_k,\dots,v_n)$ is a basis. We adjust v_k as follows: We let $a_i = \langle v_k, w_i\rangle$ and

$$w = v_k - a_1w_1 - a_2w_2 - \cdots - a_{k-1}w_{k-1}. \tag{1.24}$$

Then w is orthogonal to w_i for $i = 1,\dots,k-1$, because

$$\langle w, w_i\rangle = \langle v_k, w_i\rangle - a_1\langle w_1, w_i\rangle - a_2\langle w_2, w_i\rangle - \cdots - a_{k-1}\langle w_{k-1}, w_i\rangle.$$

Since $w_1,\dots,w_{k-1}$ are orthonormal, all the terms $\langle w_j, w_i\rangle$, $1 \le j \le k-1$, are zero except for the term $\langle w_i, w_i\rangle$, which is 1. So the sum reduces to

$$\langle w, w_i\rangle = \langle v_k, w_i\rangle - a_i\langle w_i, w_i\rangle = \langle v_k, w_i\rangle - a_i = 0.$$

We normalize the length of w to 1, obtaining a vector w_k which we substitute for v_k as before. Then $(w_1,\dots,w_k)$ is orthonormal. Since v_k is in the span of $(w_1,\dots,w_k;\ v_{k+1},\dots,v_n)$, this set is a basis. The existence of an orthonormal basis follows by induction on k. □

End of the proof of Theorem (1.19). The fact that part (iii) of Theorem (1.19) implies (i) follows from Theorem (1.22). For if A is symmetric and positive definite, then the form $\langle X,Y\rangle = X^tAY$ it defines on $\mathbb{R}^n$ is also symmetric and positive definite. In that case, Theorem (1.22) tells us that there is a basis $\mathbf{B}'$ of $\mathbb{R}^n$ which is orthonormal with respect to the form $\langle X,Y\rangle = X^tAY$. (But the basis will probably not be orthonormal with respect to the usual dot product on $\mathbb{R}^n$.) Now on the one hand, the matrix A' of the form $\langle X,Y\rangle$ with respect to the new basis $\mathbf{B}'$ satisfies the relation $P^tA'P = A$ (1.12), and on the other hand, since $\mathbf{B}'$ is orthonormal, $A' = I$. Thus $A = P^tP$. This proves (ii), and since (i) and (ii) are already known to be equivalent, it also proves (i). □

Unfortunately, there is no really simple way to show that a matrix is positive definite. One of the most convenient criteria is the following: Denote the upper left $i \times i$ submatrix of A by A_i. Thus

$$A_1 = [a_{11}],\quad A_2 = \begin{bmatrix} a_{11}a_{12} \\ a_{21}a_{22}\end{bmatrix},\quad A_3 = \begin{bmatrix} a_{11}a_{12}a_{13} \\ a_{21}a_{22}a_{23} \\ a_{31}a_{32}a_{33}\end{bmatrix},\dots, A_n = A.$$

(1.25) **Theorem.** A real symmetric $n \times n$ matrix A is positive definite if and only if the determinant $\det A_i$ is positive for each $i = 1,\dots,n$.

For example, the 2×2 matrix

$$A = \begin{bmatrix} a & b \\ b & d\end{bmatrix} \tag{1.26}$$

is positive definite if and only if $a > 0$ and $ad - bc > 0$. Using this criterion, we can check immediately that the matrix A' of (1.15) is positive definite, which agrees with the fact that it represents dot product.

The proof of Theorem (1.25) is at the end of the next section.

2. SYMMETRIC FORMS: ORTHOGONALITY

In this section, we consider a finite-dimensional real vector space V on which a symmetric bilinear form $\langle\,,\rangle$ is given, but we drop the assumption made in the last section that the form is positive definite. A form such that $\langle v, v\rangle$ takes on both positive and negative values is called *indefinite*. The *Lorentz form*

$$X^tAY = x_1y_1 + x_2y_2 + x_3y_3 - c^2x_4y_4$$

of physics is a typical example of an indefinite form on "space-time" $\mathbb{R}^4$. The coefficient c representing the speed of light can be normalized to 1, and then the matrix of the form with respect to the given basis becomes

$$(2.1) \qquad \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{bmatrix}.$$

We now pose the problem of describing all symmetric forms on a finite-dimensional real vector space. The basic notion used in the study of such a form is still that of orthogonality. But if a form is not positive definite, it may happen that a nonzero vector v is self-orthogonal: $\langle v, v\rangle = 0$. For example, this is true for the vector $(1,0,0,1)^t \in \mathbb{R}^4$, when the form is defined by (2.1). So we must revise our geometric intuition. It turns out that there is no need to worry about this point. There are enough vectors which are not self-orthogonal to serve our purposes.

(2.2) **Proposition.** Suppose the symmetric form $\langle\,,\rangle$ is not identically zero. Then there is a vector $v \in V$ which is not self-orthogonal: $\langle v, v\rangle \neq 0$.

Proof. To say that $\langle\,,\rangle$ is not identically zero means that there is a pair of vectors $v, w \in V$ such that $\langle v, w\rangle \neq 0$. Take these vectors. If $\langle v, v\rangle \neq 0$, or if $\langle w, w\rangle \neq 0$, then the proposition is verified. Suppose $\langle v, v\rangle = \langle w, w\rangle = 0$. Let $u = v + w$, and expand $\langle u, u\rangle$ using bilinearity:

$$\langle u, u\rangle = \langle v + w, v + w\rangle = \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle = 0 + 2\langle v, w\rangle + 0.$$

Since $\langle v, w\rangle \neq 0$, it follows that $\langle u, u\rangle \neq 0$. □

If W is a subspace of V, then we will denote by $W^\perp$ the set of all vectors v which are orthogonal to every $w \in W$:

$$(2.3) \qquad W^\perp = \{v \in V \mid \langle v, W\rangle = 0\}.$$

This is a subspace of V, called the *orthogonal complement* to W.

(2.4) **Proposition.** Let $w \in V$ be a vector such that $\langle w, w\rangle \neq 0$. Let $W = \{cw\}$ be the span of w. Then V is the direct sum of W and its orthogonal complement:

$$V = W \oplus W^\perp.$$

Proof. According to Chapter 3 (6.4, 6.5), we have to show two things:

(a) $W \cap W^\perp = 0$. This is clear. The vector cw is not orthogonal to w unless $c = 0$, because $\langle cw, w\rangle = c\langle w, w\rangle$ and $\langle w, w\rangle \neq 0$.

(b) W and $W^\perp$ span V: Every vector $v \in V$ can be written in the form $v = aw + v'$, where $v' \in W^\perp$. To show this, we solve the equation $\langle v - aw, w\rangle = 0$ for a: $\langle v - aw, w\rangle = \langle v, w\rangle - a\langle w, w\rangle = 0$. The solution is $a = \dfrac{\langle v, w\rangle}{\langle w, w\rangle}$. We set $v' = v - aw$. □

Two more concepts which we will need are the null space of a symmetric form and nondegenerate form. A vector $v \in V$ is called a *null vector* for the given form if $\langle v, w\rangle = 0$ for all $w \in V$, that is, if v is orthogonal to the whole space V. The *null space* of the form is the set of all null vectors

(2.5)
$$N = \{v \mid \langle v, V\rangle = 0\} = V^\perp.$$

A symmetric form is said to be *nondegenerate* if the null space is $\{0\}$.

(2.6) **Proposition.** Let A be the matrix of a symmetric form with respect to a basis.

(a) The null space of the form is the set of vectors v such that the coordinate vector X of v is a solution of the homogeneous equation $AX = 0$.
(b) The form is nondegenerate if and only if the matrix A is nonsingular.

Proof. Via the basis, the form corresponds to the product X^tAY [see (1.11)]. We might as well work with this product. If Y is a vector such that $AY = 0$, then $X^tAY = 0$ for all X; hence Y is in the null space. Conversely, suppose that $AY \neq 0$. Then AY has at least one nonzero coordinate. The ith coordinate of AY is $e_i{}^tAY$. So one of the products $e_i{}^tAY$ is not zero. This shows that Y is not a null vector, which proves (a). Part (b) of the proposition follows from (a). □

Here is a generalized version of (2.4):

(2.7) **Proposition.** Let W be a subspace of V, and consider the restriction of a symmetric form $\langle\, ,\rangle$ to W. Suppose that this form is nondegenerate on W. Then $V = W \oplus W^\perp$.

We omit the proof, which closely follows that of (2.4). □

(2.8) **Definition.** An *orthogonal basis* $\mathbf{B} = (v_1, \ldots, v_n)$ for V, with respect to a symmetric form $\langle\, ,\rangle$, is a basis such that $v_i \perp v_j$ for all $i \neq j$.

Since the matrix A of a form is defined by $a_{ij} = \langle v_i, v_j\rangle$, the basis $\mathbf{B}$ is orthogonal if and only if A is a *diagonal* matrix. Note that if the symmetric form $\langle\, ,\rangle$ is non-

degenerate and the basis $\mathbf{B} = (v_1, \ldots, v_n)$ is orthogonal, then $\langle v_i, v_i \rangle \neq 0$ for all i: the diagonal entries of A are nonzero.

(2.9) **Theorem.** Let $\langle\, , \rangle$ be a symmetric form on a real vector space V.

(a) There is an orthogonal basis for V. More precisely, there exists a basis $\mathbf{B} = (v_1, \ldots, v_n)$ such that $\langle v_i, v_j \rangle = 0$ for $i \neq j$ and such that for each i, $\langle v_i, v_i \rangle$ is either 1, -1, or 0.

(b) *Matrix form:* Let A be a real symmetric $n \times n$ matrix. There is a matrix $Q \in GL_n(\mathbb{R})$ such that QAQ^t is a diagonal matrix each of whose diagonal entries is 1, -1, or 0.

Part (b) of the theorem follows from (a), and (1.13), taking into account the fact that any symmetric matrix A is the matrix of a symmetric form. □

We can permute an orthogonal basis $\mathbf{B}$ so that the indices with $\langle v_i, v_i \rangle = 1$ are the first ones, and so on. Then the matrix A of the form will be

$$A = \begin{bmatrix} I_p & & \\ & -I_m & \\ & & 0_z \end{bmatrix}, \tag{2.10}$$

where p is the number of $+1$'s, m is the number of -1's, and z is the number of 0's, so that $p + m + z = n$. These numbers are uniquely determined by the form or by the matrix A:

(2.11) **Theorem.** *Sylvester's Law:* The numbers p, m, z appearing in (2.10) are uniquely determined by the form. In other words, they do not depend on the choice of orthogonal basis $\mathbf{B}$ such that $\langle v_i, v_i \rangle = \pm 1$ or 0.

The pair of integers (p, m) is called the *signature* of the form.

Proof of Theorem (2.9). If the form is identically zero, then the matrix A, computed with respect to any basis, will be the zero matrix, which is diagonal. Suppose the form is not identically zero. Then by Proposition (2.2), there is a vector $v = v_1$ with $\langle v_1, v_1 \rangle \neq 0$. Let W be the span of v_1. By Proposition (2.4), $V = W \oplus W^{\perp}$, and so a basis for V is obtained by combining the basis (v_1) of W with any basis $(v_2, \ldots, v_n)$ of $W^{\perp}$ [Chapter 3 (6.6)]. The form on V can be restricted to the subspace $W^{\perp}$, and it defines a form there. We use induction on the dimension to conclude that $W^{\perp}$ has an orthogonal basis $(v_2, \ldots, v_n)$. Then $(v_1, v_2, \ldots, v_n)$ is an orthogonal basis for V. For, $\langle v_1, v_i \rangle = 0$ if $i > 1$ because $v_i \in W^{\perp}$, and $\langle v_i, v_j \rangle = 0$ if $i, j > 1$ and $i \neq j$, because $(v_2, \ldots, v_n)$ is an orthogonal basis.

It remains to normalize the orthogonal basis just constructed. If $\langle v_i, v_i \rangle \neq 0$, we solve $c^{-2} = \pm \langle v_i, v_i \rangle$ and change the basis vector v_i to cv_i. Then $\langle v_i, v_i \rangle$ is changed to ± 1. This completes the proof of (2.9.) □

Proof of Theorem (2.11). Let $r = p + m$. (This is the rank of the matrix A.) Let $(v_1,\ldots,v_n)$ be an orthogonal basis of V of the type under consideration, that is, so that the matrix is (2.10). We will first show that the number z is determined by proving that the vectors $v_{r+1},\ldots,v_n$ form a basis for the null space $N = V^{\perp}$. This will show that $z = \dim N$, hence that z does not depend on the choice of a basis.

A vector $w \in V$ is a null vector if and only if it is orthogonal to every element v_i of our basis. We write our vector as a linear combination of the basis: $w = c_1v_1 + \cdots + c_nv_n$. Then since $\langle v_i, v_j\rangle = 0$ if $i \neq j$, we find $\langle w, v_i\rangle = c_i\langle v_i, v_i\rangle$. Now $\langle v_i, v_i\rangle = 0$ if and only if $i > r$. So in order for w to be orthogonal to every v_i, we must have $c_i = 0$ for all $i \leq r$. This shows that $(v_{r+1},\ldots,v_n)$ spans N, and, being a linearly independent set, it is a basis for N.

The equation $p + m + z = n$ proves that $p + m$ is also determined. We still have to show that one of the two remaining integers p,m is determined. This is not quite so simple. It is not true that the span of $(v_1,\ldots,v_p)$, for instance, is uniquely determined by the form.

Suppose a second such basis $(v_1',\ldots,v_n')$ is given and leads to integers p',m' (with $z' = z$). We will show that the $p + (n - p')$ vectors

$$v_1,\ldots,v_p;\ v_{p'+1}',\ldots,v_n' \tag{2.12}$$

are linearly independent. Then since V has dimension n, it will follow that $p + (n - p') \leq n$, hence that $p \leq p'$, and, interchanging the roles of p and p', that $p = p'$.

Let a linear relation between the vectors (2.12) be given. We may write it in the form

$$b_1v_1 + \cdots + b_pv_p = c_{p'+1}v_{p'+1}' + \cdots + c_nv_n'. \tag{2.13}$$

Let v denote the vector defined by either of these two expressions. We compute $\langle v, v\rangle$ in two ways. The left-hand side gives

$$\langle v, v\rangle = b_1^2\langle v_1, v_1\rangle + \cdots + b_p^2\langle v_p, v_p\rangle = b_1^2 + \cdots + b_p^2 \geq 0,$$

while the right-hand side gives

$$\langle v, v\rangle = c_{p'+1}^2\langle v_{p'+1}, v_{p'+1}\rangle + \cdots + c_n^2\langle v_n', v_n'\rangle = -c_{p'+1}^2 - \cdots - c_{p'+m'}^2 \leq 0.$$

It follows that $b_1^2 + \cdots + b_p^2 = 0$, hence that $b_1 = \ldots = b_p = 0$. Once this is known, the fact that $(v_1',\ldots,v_n')$ is a basis combines with (2.13) to imply $c_{p'+1} = \ldots = c_n = 0$. Therefore the relation was trivial, as required. □

For dealing with indefinite forms, the notation $I_{p,m}$ is often used to denote the diagonal matrix

$$I_{p,m} = \begin{bmatrix} I_p & \\ & -I_m \end{bmatrix}. \tag{2.14}$$

With this notation, the matrix representing the Lorentz form (2.1) is $I_{3,1}$.

We will now prove Theorem (1.25)—that a matrix A is positive definite if and only if $\det A_i > 0$ for all i.

Proof of Theorem (1.25). Suppose that the form X^tAY is positive definite. A change of basis in $\mathbb{R}^n$ changes the matrix to $A' = QAQ^t$, and

$$\det A' = (\det Q)(\det A)(\det Q^t) = (\det Q)^2(\det A).$$

Since they differ by a square factor, $\det A'$ is positive if and only if $\det A$ is positive. By (1.19), we can choose a matrix Q so that $A' = I$, and since I has determinant 1, $\det A > 0$.

The matrix A_i represents the restriction of the form to the subspace V_i spanned by $(v_1, \ldots, v_i)$, and of course the form is positive definite on V_i. Therefore $\det A_i > 0$ for the same reason that $\det A > 0$.

Conversely, suppose that $\det A_i$ is positive for all i. By induction on n, we may assume the form to be positive definite on V_{n-1}. Therefore there is a matrix $Q' \in GL_{n-1}$ such that $Q'A_{n-1}Q'^t = I_{n-1}$. Let Q be the matrix

$$Q = \begin{bmatrix} Q' & \\ & 1 \end{bmatrix}.$$

Then

$$QAQ^t = \begin{bmatrix} & & * \\ & I & \vdots \\ * & \cdots & * \end{bmatrix}.$$

We now clear out the bottom row of this matrix, except for the (n,n) entry, by elementary row operations $E_1, \ldots, E_{n-1}$. Let $P = E_{n-1} \cdots E_1 Q$. Then

$$A' = PAP^t = \begin{bmatrix} \begin{bmatrix} & & \\ & I & \\ & & \end{bmatrix} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ 0 \quad \cdots \quad 0 & c \end{bmatrix},$$

for some c. The last column has also been cleared out because $A' = PAP^t$ is symmetric. Since $\det A > 0$, we have $\det A' = (\det A)(\det P)^2 > 0$ too, and this implies that $c > 0$. Therefore the matrix A' represents a positive definite form. It also represents the same form as A does. So A is positive definite. □

3. THE GEOMETRY ASSOCIATED TO A POSITIVE FORM

In this section we return to look once more at a positive definite bilinear form $\langle\ ,\ \rangle$ on an n-dimensional real vector space V. A real vector space together with such a form is often called a *Euclidean space*.

It is natural to define the *length* of a vector v by the rule

$$|v| = \sqrt{\langle v, v \rangle}, \tag{3.1}$$

in analogy with the length of vectors in $\mathbb{R}^n$ [Chapter 4 (5.10)]. One important consequence of the fact that the form is positive definite is that we can decide whether a

vector v is zero by computing its length:

$$v = 0 \quad \textit{if and only if} \quad \langle v, v\rangle = 0. \tag{3.2}$$

As was shown in Section 1, there is an orthonormal basis $\mathbf{B} = (v_1, \ldots, v_n)$ for V, and thereby the form corresponds to dot product on $\mathbb{R}^n$:

$$\langle v, w\rangle \quad = \quad X^tY,$$

if $v = \mathbf{B}X$ and $w = \mathbf{B}Y$. Using this correspondence, we can transfer the geometry of $\mathbb{R}^n$ over to V. Whenever a problem is presented to us on a Euclidean space V, a natural procedure will be to choose a convenient orthonormal basis, thereby reducing the problem to the familiar case of dot product on $\mathbb{R}^n$.

When a subspace W of V is given to us, there are two operations we can make. The first is to *restrict* the form $\langle\, ,\rangle$ to the subspace, simply by defining the value of the form on a pair w_1, w_2 of vectors in W to be $\langle w_1, w_2\rangle$. The restriction of a bilinear form to a subspace W is a bilinear form on W, and if the form is symmetric or if it is symmetric and positive definite, then so is the restriction.

Restriction of the form can be used to define the unoriented *angle* between two vectors v, w. If the vectors are linearly dependent, the angle is zero. Otherwise, (v, w) is a basis of a two-dimensional subspace W of V. The restriction of the form to W is still positive definite, and therefore there is an orthonormal basis (w_1, w_2) for W. By means of this basis, v, w correspond to their coordinate vectors X, Y in $\mathbb{R}^2$. This allows us to interpret geometric properties of the vectors v, w in terms of properties of X, Y.

Since the basis (w_1, w_2) is orthonormal, the form corresponds to dot product on $\mathbb{R}^2$: $\langle v, w\rangle = X^tY$. Therefore

$$|v| = |X|, \quad |w| = |Y|, \quad \text{and} \quad \langle v, w\rangle = (X \cdot Y).$$

We define the angle θ between v and w to be the angle between X and Y, and thereby obtain the formula

$$\langle v, w\rangle = |v||w| \cos \theta, \tag{3.3}$$

as a consequence of the analogous formula [Chapter 4 (5.11)] for dot product in $\mathbb{R}^2$. This formula determines $\cos\theta$ in terms of the other symbols, and $\cos\theta$ determines θ up to a factor of ± 1. Therefore the angle between v and w is determined up to sign by the form alone. This is the best that can be done, even in $\mathbb{R}^3$.

Standard facts such as the *Schwarz Inequality*

$$|\langle v, w\rangle| \leq |v||w| \tag{3.4}$$

and the *Triangle Inequality*

$$|v + w| \leq |v| + |w| \tag{3.5}$$

can also be proved for arbitrary Euclidean spaces by restriction to a two-dimensional subspace.

The second operation we can make when a subspace W is given is to project V onto W. Since the restriction of the form to W is positive definite, it is nondegenerate. Therefore $V = W \oplus W^{\perp}$ by (2.17), and so every $v \in V$ has a unique expression

(3.6) $$v = w + w', \quad \text{with} \quad w \in W \quad \text{and} \quad \langle w, w' \rangle = 0.$$

The *orthogonal projection* $\pi: V \longrightarrow W$ is defined to be the linear transformation

(3.7) $$v \rightsquigarrow \pi(v) = w$$

where w is as in (3.6).

The projected vector $\pi(v)$ can be computed easily in terms of an orthonormal basis $(w_1, \ldots, w_r)$ of W. What follows is important:

(3.8) **Proposition.** Let $(w_1, \ldots, w_r)$ be an orthonormal basis of a subspace W, and let $v \in V$. The orthogonal projection $\pi(v)$ of v onto w is the vector

$$\pi(v) = \langle v, w_1 \rangle w_1 + \cdots + \langle v, w_r \rangle w_r.$$

Thus if π is defined by the above formula, then $v - \pi(v)$ is orthogonal to W. This formula explains the geometric meaning of the Gram–Schmidt procedure described in Section 1.

Proof. Let us denote the right side of the above equation by $\tilde{w}$. Then $\langle \tilde{w}, w_i \rangle = \langle v, w_i \rangle \langle w_i, w_i \rangle = \langle v, w_i \rangle$ for $i = 1, \ldots, r$, hence $v - \tilde{w} \in W^{\perp}$. Since the expression (3.6) for v is unique. $w = \tilde{w}$ and $w' = v - \tilde{w}$. □

The case $W = V$ is also important. In this case, π is the identity map.

(3.9) **Corollary.** Let $\mathbf{B} = (v_1, \ldots, v_n)$ be an orthonormal basis for a Euclidean space V. Then

$$v = \langle v, v_1 \rangle v_1 + \cdots + \langle v, v_n \rangle v_n.$$

In other words, the coordinate vector of v with respect to the orthonormal basis $\mathbf{B}$ is

$$X = (\langle v, v_1 \rangle, \ldots, \langle v, v_n \rangle)^t. \;\square$$

4. HERMITIAN FORMS

In this section we assume that our scalar field is the field $\mathbb{C}$ of complex numbers. When working with complex vector spaces, it is desirable to have an analogue of the concept of the length of a vector, and of course one can define length on $\mathbb{C}^n$ by identifying it with $\mathbb{R}^{2n}$. If $X = (x_1, \ldots, x_n)^t$ is a complex vector and if $x_r = a_r + b_r i$, then the *length* of X is

(4.1) $$|X| = \sqrt{a_1^2 + b_1^2 + \cdots + a_n^2 + b_n^2} = \sqrt{\bar{x}_1 x_1 + \cdots + \bar{x}_n x_n},$$

where the bar denotes complex conjugation. This formula suggests that dot product

is "wrong" for complex vectors and that we should define a product by the formula

$$\langle X, Y\rangle = \overline{X}^{t}Y = \overline{x}_1 y_1 + \cdots + \overline{x}_n y_n. \tag{4.2}$$

This product has the *positivity* property:

(4.3) *$\langle X, X\rangle$ is a positive real number if $X \neq 0$.*

Moreover, (4.2) agrees with dot product for real vectors.

The product (4.2) is called the *standard hermitian product*, or the *hermitian dot product*. It has these properties:

(4.4)

Linearity in the second variable:

$$\langle X, cY\rangle = c\langle X, Y\rangle \quad \text{and} \quad \langle X, Y_1 + Y_2\rangle = \langle X, Y_1\rangle + \langle X, Y_2\rangle;$$

Conjugate linearity in the first variable:

$$\langle cX, Y\rangle = \overline{c}\langle X, Y\rangle \quad \text{and} \quad \langle X_1 + X_2, Y\rangle = \langle X_1, Y\rangle + \langle X_2, Y\rangle;$$

Hermitian symmetry:

$$\langle Y, X\rangle = \overline{\langle X, Y\rangle}.$$

So we can have a positive definite product at a small cost in linearity and symmetry.

When one wants to work with notions involving length, the hermitian product is the right one, though symmetric bilinear forms on complex vector spaces also come up in applications.

If V is a complex vector space, a *hermitian form* on V is any function of two variables

$$\begin{aligned} V \times V &\longrightarrow \mathbb{C} \\ v, w &\rightsquigarrow \langle v, w\rangle \end{aligned} \tag{4.5}$$

satisfying the relations (4.4). Let $\mathbf{B} = (v_1, \ldots, v_n)$ be a basis for V. Then the *matrix* of the form is defined in the analogous way as the matrix of a bilinear form:

$$A = (a_{ij}), \quad \text{where } a_{ij} = \langle v_i, v_j\rangle.$$

The formula for the form now becomes

$$\langle v, w\rangle = \overline{X}^{t}AY, \tag{4.6}$$

if $v = \mathbf{B}X$ and $w = \mathbf{B}Y$.

The matrix A is not arbitrary, because hermitian symmetry implies that

$$a_{ij} = \langle v_i, v_j\rangle = \overline{\langle v_j, v_i\rangle} = \overline{a}_{ji},$$

that is, that $A = \overline{A}^{t}$. Let us introduce the *adjoint* of a matrix A [different from the

one defined in Chapter 1 (5.4)] as

(4.7) $$A^* = \overline{A}^t.$$

It satisfies the following rules:

$$(A+B)^* = A^* + B^*$$
$$(AB)^* = B^*A^*$$
$$(A^*)^{-1} = (A^{-1})^*$$
$$A^{**} = A.$$

These rules are easy to check. Formula (4.6) can now be rewritten as

(4.8) $$\langle v, w\rangle = X^*AY,$$

and the standard hermitian product on $\mathbb{C}^n$ becomes $\langle X, Y\rangle = X^*Y$.

A matrix A is called *hermitian* or *self-adjoint* if

(4.9) $$A = A^*,$$

and it is the hermitian matrices which are matrices of hermitian forms. Their entries satisfy $a_{ji} = \overline{a}_{ij}$. This implies that the diagonal entries are real and that the entries below the diagonal are complex conjugates of those above it:

$$A = \begin{bmatrix} r_1 & & a_{ij} \\ & \ddots & \\ \overline{a}_{ij} & & r_n \end{bmatrix}, \quad r_i \in \mathbb{R}, \quad a_{ij} \in \mathbb{C}.$$

For example, $\begin{bmatrix} 2 & i \\ -i & 1 \end{bmatrix}$ is a hermitian matrix.

Note that the condition for a real matrix to be hermitian is $a_{ji} = a_{ij}$:

(4.10) *The real hermitian matrices are the real symmetric matrices.*

The discussion of change of basis in Sections 1 and 2 has analogues for hermitian forms. Given a hermitian form, a change of basis by a matrix P leads as in (1.12) to

$$X'^*A'Y' = (PX)^*A'PY = X^*(P^*A'P)Y.$$

Hence the new matrix A' satisfies

(4.11) $$A = P^*A'P \quad \text{or} \quad A' = (P^*)^{-1}AP^{-1}.$$

Since P is arbitrary, we can replace it by $Q = (P^*)^{-1}$ to obtain the description analogous to (1.13):

(4.12) **Corollary.** Let A be the matrix of a hermitian form with respect to a basis. The matrices which represent the same hermitian form with respect to different bases are those of the form $A' = QAQ^*$, for some invertible matrix $Q \in GL_n(\mathbb{C})$. □

For hermitian forms, the analogues of orthogonal matrices are the unitary matrices. A matrix P is called *unitary* if it satisfies the condition

$$(4.13) \qquad P^*P = I \quad \text{or} \quad P^* = P^{-1}.$$

For example, $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix}$ is a unitary matrix.

Note that for a real matrix P, this condition becomes $P^tP = I$:

(4.14) *The real unitary matrices are the real orthogonal matrices.*

The unitary matrices form a group, the *unitary group* U_n:

$$(4.15) \qquad U_n = \{P \mid P^*P = I\}.$$

Formula (4.11) tells us that unitary matrices represent changes of basis which leave the standard hermitian product X^*Y invariant:

(4.16) **Corollary.** A change of basis preserves the standard hermitian product, that is, $X^*Y = X'^*Y'$, if and only if its matrix P is unitary. □

But Corollary (4.12) tells us that a general change of basis changes the standard hermitian product X^*Y to $X'^*A'Y'$, where $A' = QQ^*$, and $Q \in GL_n(\mathbb{C})$.

The notion of orthogonality for hermitian forms is defined exactly as for symmetric bilinear forms: v is called *orthogonal* to w if $\langle v, w \rangle = 0$. Since $\overline{\langle v, w \rangle} = \langle w, v \rangle$, orthogonality is still a symmetric relation. We can now copy the discussion of Sections 1 and 2 for hermitian forms without essential change, and Sylvester's Law (2.11) for real symmetric forms carries over to the hermitian case. In particular, we can speak of *positive definite* forms, those having the property that

(4.17) $\langle v, v \rangle$ *is a positive real number if* $v \neq 0$,

and of *orthonormal bases* $\mathbf{B} = (v_1, \ldots, v_n)$, those such that

$$(4.18) \qquad \langle v_i, v_i \rangle = 1 \quad \text{and} \quad \langle v_i, v_j \rangle = 0 \quad \text{if } i \neq j.$$

(4.19) **Theorem.** Let $\langle\, ,\rangle$ be a hermitian form on a complex vector space V. There is an orthonormal basis for V if and only if the form is positive definite.

(4.20) **Proposition.** Let W be a subspace of a hermitian space V. If the restriction of the form to W is nondegenerate, then $V = W \oplus W^{\perp}$

The proofs of these facts are left as exercises. □

5. THE SPECTRAL THEOREM

In this section we will study an n-dimensional complex vector space V and a positive definite hermitian form $\langle\, ,\rangle$ on V. A complex vector space on which a positive definite hermitian form is given is often called a *hermitian space*. You can imagine that V is $\mathbb{C}^n$, with its standard hermitian product X^*Y, if you want to. The choice of an orthonormal basis in V will allow such an identification.

Since the form $\langle\, ,\rangle$ is given, we will not want to choose an arbitrary basis for V in order to make computations. It is natural to work exclusively with orthonormal bases. This changes all previous calculations in the following way: It will no longer be true that the matrix P of a change of basis is an arbitrary invertible matrix. Rather, if $\mathbf{B} = (v_1,\dots,v_n)$, $\mathbf{B}' = (v_1',\dots,v_n')$ are two *orthonormal* bases, then the matrix P relating them will be unitary. The fact that the bases are orthonormal means that the matrix of the form $\langle\, ,\rangle$ with respect to each basis is the identity I, and so (4.11) reads $I = P^*IP$, or $P^*P = I$.

We are going to study a linear operator

$$T: V \longrightarrow V \tag{5.1}$$

on our space. Let $\mathbf{B}$ be an orthonormal basis, and let M be the associated matrix of T. A change of orthonormal basis changes M to $M' = PMP^{-1}$ [Chapter 4 (3.4)] where P is unitary; hence

$$M' = PMP^*. \tag{5.2}$$

(5.3) **Proposition.** Let T be a linear operator on a hermitian space V, and let M be the matrix of T with respect to an orthonormal basis $\mathbf{B}$.

(a) The matrix M is hermitian if and only if $\langle v, Tw\rangle = \langle Tv, w\rangle$ for all $v, w \in V$. If so, T is called a *hermitian operator*.

(b) The matrix M is unitary if and only if $\langle v, w\rangle = \langle Tv, Tw\rangle$ for all $v, w \in V$. If so, T is called a *unitary operator*.

Proof. Let X, Y be the coordinate vectors of v, w: $v = \mathbf{B}X$, $w = \mathbf{B}Y$, so that $\langle v, w\rangle = X^*Y$ and $Tv = \mathbf{B}MX$. Then $\langle v, Tw\rangle = X^*MY$, and $\langle Tv, w\rangle = X^*M^*Y$. So if $M = M^*$, then $\langle v, Tw\rangle = \langle Tv, w\rangle$ for all v,w; that is, T is hermitian. Conversely, if T is hermitian, we set $v = e_i$, $w = e_j$ as in the proof of (1.9) to obtain $b_{ij} = e_i^*(Me_j) = (e_i^*M^*)e_j = \bar{b}_{ji}$. Thus $M = M^*$. Similarly, $\langle v, w\rangle = X^*Y$ and $\langle Tv, Tw\rangle = X^*M^*MY$, so $\langle v, w\rangle = \langle Tv, Tw\rangle$ for all v, w if and only if $M^*M = I$. □

(5.4) **Theorem.** *Spectral Theorem:*

(a) Let T be a hermitian operator on a hermitian vector space V. There is an orthonormal basis of V consisting of eigenvectors of T.

(b) *Matrix form:* Let M be a hermitian matrix. There is a unitary matrix P such that PMP^* is a real diagonal matrix.

Proof. Choose an eigenvector $v = v_1$, and normalize so that its length is 1: $\langle v, v\rangle = 1$. Extend to an orthonormal basis. Then the matrix of T becomes

$$M = \begin{bmatrix} a & * & \cdots & * \\ 0 & & & \\ \vdots & & N & \\ 0 & & & \end{bmatrix}.$$

Since T is hermitian, so is the matrix M (5.3). This implies that $* \cdots * = 0 \cdots 0$ and that N is hermitian. Proceed by induction. □

To diagonalize a hermitian matrix M by a unitary P, one can proceed by determining the eigenvectors. If the eigenvalues are distinct, the corresponding eigenvectors will be orthogonal. This follows from the Spectral Theorem. Let $\mathbf{B}'$ be the orthonormal basis obtained by normalizing the lengths of the eigenvectors to 1. Then $P = [\mathbf{B}']^{-1}$ [Chapter 3 (4.20)].

For example, let

$$M = \begin{bmatrix} 2 & i \\ -i & 2 \end{bmatrix}.$$

The eigenvalues of this matrix are 3, 1, and the vectors

$$v_1' = \begin{bmatrix} 1 \\ -i \end{bmatrix}, \quad v_2' = \begin{bmatrix} 1 \\ i \end{bmatrix}$$

are eigenvectors with these eigenvalues. We normalize their lengths to 1 by the factor $\frac{1}{\sqrt{2}}$. Then

$$P = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix}^* = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \quad \text{and} \quad PMP^* = \begin{bmatrix} 3 & \\ & 1 \end{bmatrix}.$$

But the Spectral Theorem asserts that a hermitian matrix can be diagonalized *even if its eigenvalues aren't distinct*. This statement becomes particularly simple for 2×2 matrices: If the characteristic polynomial of a 2×2 hermitian matrix M has a double root, then there is a unitary matrix P such that $PMP^* = aI$. Bringing the P's over to the other side of the equation, we obtain $M = P^*aIP = aP^*P = aI$. So it follows from the Spectral Theorem that $M = aI$. The only 2×2 hermitian matrices whose characteristic polynomials have a double root are the matrices aI, where a is a real number. We can verify this fact directly from the definition. We write $M = \begin{bmatrix} a & \beta \\ \overline{\beta} & d \end{bmatrix}$, where a, d are real and β is complex. Then the characteristic polynomial is $t^2 - (a + d)t + (ad - \beta\overline{\beta})$. This polynomial has a double root if and only if its discriminant vanishes, that is, if

$$(a + d)^2 - 4(ad - \beta\overline{\beta}) = (a - d)^2 + 4\beta\overline{\beta} = 0.$$

Both of the terms $(a - d)^2$ and $\beta\overline{\beta}$ are nonnegative real numbers. So if the discriminant vanishes, then $a = d$ and $\beta = 0$. In this case, $M = aI$, as predicted.

Here is an interesting consequence of the Spectral Theorem for which we can give a direct proof:

(5.5) **Proposition.** The eigenvalues of a hermitian operator T are real numbers.

Proof. Let a be an eigenvalue, and let v be an eigenvector for T such that $T(v) = av$. Then by (5.3) $\langle Tv, v\rangle = \langle v, Tv\rangle$; hence $\langle av, v\rangle = \langle v, av\rangle$. By conjugate linearity (4.4),

$$\overline{a}\langle v, v\rangle = \langle av, v\rangle = \langle v, av\rangle = a\langle v, v\rangle,$$

and $\langle v, v\rangle \neq 0$ because the form $\langle\, ,\rangle$ is positive definite. Hence $a = \overline{a}$. This shows that a is real. □

The results we have proved for hermitian matrices have analogues for real symmetric matrices. Let V be a real vector space with a positive definite bilinear form $\langle\, ,\rangle$. Let T be a linear operator on V.

(5.6) **Proposition.** Let M be the matrix of T with respect to an orthonormal basis.

(a) The matrix M is symmetric if and only if $\langle v, Tw\rangle = \langle Tv, w\rangle$ for all $v, w \in V$. If so, T is called a *symmetric operator*.

(b) The matrix M is orthogonal if and only if $\langle v, w\rangle = \langle Tv, Tw\rangle$ for all $v, w \in V$. If so, T is called an *orthogonal operator*. □

(5.7) **Proposition.** The eigenvalues of a real symmetric matrix are real.

Proof. A real symmetric matrix is hermitian. So this is a special case of (5.5). □

(5.8) **Theorem.** *Spectral Theorem (real case):*

(a) Let T be a symmetric operator on a real vector space V with a positive definite bilinear form. There is an orthonormal basis of eigenvectors of T.

(b) *Matrix form:* Let M be a real symmetric $n \times n$ matrix. There is an orthogonal matrix $P \in O_n(\mathbb{R})$ such that PMP^t is diagonal.

Proof. Now that we know that the eigenvalues of such an operator are real, we can copy the proof of (5.4). □

6. CONICS AND QUADRICS

A *conic* is the locus in the plane $\mathbb{R}^2$ defined by a quadratic equation in two variables, of the form

$$f(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + b_1x_1 + b_2x_2 + c = 0. \tag{6.1}$$

More precisely, the locus (6.1) is a conic, meaning an ellipse, a hyperbola, or a parabola, or else it is called *degenerate*. A degenerate conic can be a pair of lines, a single line, a point, or empty, depending on the particular equation. The term *quadric* is used to designate the analogous loci in three or more dimensions.

The quadratic part of $f(x_1, x_2)$ is called a quadratic form:

$$q(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2. \tag{6.2}$$

In general, a *quadratic form* in n variables $x_1,\dots,x_n$ is a polynomial each of whose terms has degree 2 in the variables.

It is convenient to express the form $q(x_1, x_2)$ in matrix notation. To do this, we introduce the symmetric matrix

$$A = \begin{bmatrix} a_{11} a_{12} \\ a_{12} a_{22} \end{bmatrix}. \tag{6.3}$$

Then $q(x_1, x_2) = X^tAX$, where X denotes the column vector $(x_1, x_2)^t$. We also introduce the row vector $B = (b_1, b_2)$. Then equation (6.1) can be written in matrix notation as

$$X^tAX + BX + c = 0. \tag{6.4}$$

We put the coefficient 2 into formulas (6.1) and (6.2) in order to avoid some coefficients $\frac{1}{2}$ in the matrix (6.3). An alternative way to write the quadratic form would be

$$q(x_1, x_2) = a_{11}x_1^2 + a_{12}x_1x_2 + a_{12}x_2x_1 + a_{22}x_2^2.$$

We propose to describe the congruence classes of conics as geometric figures or, what is the same, their orbits under the action of the group M of rigid motions of the plane. A rigid motion will produce a change of variable in equation (6.1).

(6.5) **Theorem.** Every nondegenerate conic is congruent to one of the following:

(i) *Ellipse:* $a_{11}x_1^2 + a_{22}x_2^2 - 1 = 0,$

(ii) *Hyperbola:* $a_{11}x_1^2 - a_{22}x_2^2 - 1 = 0,$

(iii) *Parabola:* $a_{11}x_1^2 - x_2 = 0,$ where $a_{11}, a_{22} > 0$.

Proof. We simplify equation (6.1) in two steps, first applying an orthogonal transformation (a rotation or reflection) to diagonalize A and then applying a translation to eliminate, as much as possible, the linear and constant terms $BX + c$.

By the Spectral Theorem (5.8), there is an orthogonal matrix P such that PAP^t is diagonal. We make the change of variable $X' = PX$, or $X = P^tX'$. Substitution into equation (6.4) yields

$$X'^t(PAP^t)X' + (BP^t)X' + c = 0. \tag{6.6}$$

Hence there is an orthogonal change of variable such that the quadratic form becomes diagonal, that is, the coefficient a_{12} of x_1x_2 is zero.

Suppose that A is diagonal. Then f has the form

$$f(x_1, x_2) = a_{11}x_1^2 + a_{22}x_2^2 + b_1x_1 + b_2x_2 + c = 0.$$

We eliminate b_i by completing the squares, making the substitution

$$x_i = \left(x_i' - \frac{b_i}{2a_{ii}}\right). \tag{6.7}$$

This substitution results in

$$f(x_1, x_2) = a_{11}x_1'^2 + a_{22}x_2'^2 + c', \tag{6.8}$$

where c' is a number which can be determined if desired. This substitution corresponds to translation by the vector $(b_1/2a_{11}, b_2/2a_{22})^t$, and we can make it provided a_{11}, a_{22} are not zero.

If $a_{ii} = 0$ but $b_i \neq 0$, then we can use the substitution

$$x_i = x_i' - c/b_i \tag{6.9}$$

to eliminate the constant term instead. We may normalize one coefficient to -1. Doing so and eliminating degenerate conics leaves us with the three cases listed in the theorem. It is not difficult to show that a change of the coefficients a_{11}, a_{22} results in a different congruence class, except for the interchange of a_{11}, a_{22} in the equation of an ellipse. □

The method used above can be applied in any number of variables to classify quadrics in n dimensions. The general quadratic equation has the form

$$f(x_1, \ldots, x_n) = \sum_i a_{ii}x_i^2 + \sum_{i<j} 2a_{ij}x_ix_j + \sum_i b_ix_i + c = 0. \tag{6.10}$$

We could also write this equation more compactly as

$$f(x_1, \ldots, x_n) = \sum_{i,j} a_{ij}x_ix_j + \sum_i b_ix_i + c = 0, \tag{6.11}$$

where the first sum is over all pairs of indices, and where we set $a_{ji} = a_{ij}$.

We define the matrices A,B to be

$$A = \begin{bmatrix} a_{11}a_{12} & \cdots & a_{1m} \\ a_{12} & & \vdots \\ \vdots & & \vdots \\ a_{1m} & \cdots & a_{mm} \end{bmatrix}, \quad B = (b_1, \ldots, b_n).$$

Then the quadratic form is

$$q(x_1, \ldots, x_n) = X^tAX, \tag{6.12}$$

and

$$f(x_1, \ldots, x_n) = X^tAX + BX + c. \tag{6.13}$$

By a suitable orthogonal transformation P, the quadric is carried to (6.6), where PAP^t is diagonal. When A is diagonal, linear terms are eliminated by the translation (6.7), or else (6.9) is used.

Here is the classification in three variables:

(6.14) **Theorem.** The congruence classes of nondegenerate quadrics in $\mathbb{R}^3$ are represented by

(i) *Ellipsoids:* $a_{11}x_1^2+a_{22}x_2^2+a_{33}x_3^2-1 = 0,$

(ii) *1-sheeted hyperboloids:* $a_{11}x_1^2+a_{22}x_2^2-a_{33}x_3^2-1 = 0,$

(iii) *2-sheeted hyperboloids:* $a_{11}x_1^2-a_{22}x_2^2-a_{33}x_3^2-1 = 0,$

(iv) *Elliptic paraboloids:* $a_{11}x_1^2+a_{22}x_2^2-x_3 = 0,$

(v) *Hyperbolic paraboloids:* $a_{11}x_1^2-a_{22}x_2^2-x_3 = 0,$

where $a_{11}, a_{22}, a_{33} > 0$. □

If a quadratic equation $f(x_1, x_2) = 0$ is given, we can determine the type of conic it represents most easily by allowing nonorthogonal changes of coordinates. For example, if the associated quadratic form q is positive definite, then the conic is either an ellipse, or else it is degenerate (a single point or empty). To distinguish these cases, arbitrary changes of coordinates are permissible. A nonorthogonal coordinate change will distort the conic, but it will not change an ellipse into a hyperbola or a degenerate conic.

As an example, consider the locus

$$x_1^2+x_1x_2+x_2^2 + 4x_1+3x_2 + 4 = 0. \tag{6.15}$$

The associated matrix is

$$A = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix},$$

which is positive definite by (1.25). We diagonalize A by the nonorthogonal substitution $X' = PX$, where

$$P = \begin{bmatrix} 1 & \frac{1}{2} \\ & 1 \end{bmatrix}, \quad PAP^t = \begin{bmatrix} 1 & \\ & \frac{3}{4} \end{bmatrix}, \quad BP^t = (4, 1),$$

to obtain

$$x_1'^2+\tfrac{3}{4}x_2'^2 + 4x_1'+x_2' + 4 = 0.$$

Completing the square yields

$$x_1''^2 + \tfrac{3}{4}x_2''^2 - \tfrac{1}{3} = 0,$$

an ellipse. Thus (6.15) represents an ellipse too. On the other hand, if we change the constant term of (6.15) to 5, the locus becomes empty.

7. THE SPECTRAL THEOREM FOR NORMAL OPERATORS

The Spectral Theorem (5.4) tells us that any hermitian matrix M can be transformed into a real diagonal matrix D by a unitary matrix P: $D = PMP^*$. We now ask for the matrices M which can be transformed in the same way to a diagonal matrix D, but where we no longer require D to be real. It turns out that there is an elegant formal characterization of such matrices.

(7.1) **Definition.** A matrix M is called *normal* if it commutes with its adjoint, that is, if $MM^* = M^*M$.

(7.2) **Lemma.** If M is normal and P is unitary, then $M' = PMP^*$ is also normal, and conversely.

Proof. Assume that M is normal. Then $M'M'^* = PMP^*(PMP^*)^* = PMM^*P^* = PM^*MP^* = (PMP^*)^*(PMP^*) = M'^*M'$. So PMP^* is normal. The converse follows by replacing P by P^*. □

This lemma allows us to define a *normal operator* $T: V \longrightarrow V$ on a hermitian space V to be a linear operator whose matrix M with respect to any orthonormal basis is a normal matrix.

(7.3) **Theorem.** A complex matrix M is normal if and only if there is a unitary matrix P such that PMP^* is diagonal. □

The most important normal matrices, aside from hermitian ones, are unitary matrices: Since $M^* = M^{-1}$ if M is unitary, $MM^* = M^*M = I$, which shows that M is normal.

(7.4) **Corollary.** Every conjugacy class in the unitary group contains a diagonal matrix. □

Proof of Theorem (7.3). First, any two diagonal matrices commute, so a diagonal matrix is normal: $DD^* = D^*D$. The lemma tells us that M is normal if $PMP^* = D$. Conversely, suppose that M is normal. Choose an eigenvector $v = v_1$ of M, and normalize so that $\langle v, v \rangle = 1$, as in the proof of (5.4). Extend $\{v_1\}$ to an orthonormal basis. Then M will be changed to a matrix

$$M_1 = PMP^* = \begin{bmatrix} a_{11} a_{12} \cdots a_{1n} \\ \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \begin{bmatrix} & & \\ & N & \\ & & \end{bmatrix} \end{bmatrix}, \quad \text{and} \quad M_1^* = PM^*P^* = \begin{bmatrix} a_{11} \; 0 \; \cdots \; 0 \\ \begin{matrix} \overline{a}_{12} \\ \vdots \\ \overline{a}_{1n} \end{matrix} \begin{bmatrix} & & \\ & N^* & \\ & & \end{bmatrix} \end{bmatrix}.$$

The upper left entry of $M_1^*M_1$ is $a_{11}\overline{a}_{11}$, while the same entry of $M_1M_1^*$ is $a_{11}\overline{a}_{11}+a_{12}\overline{a}_{12}+\cdots+a_{1n}\overline{a}_{1n}$. Since M is normal, so is M_1, that is, $M_1^*M_1^* = M_1M_1^*$. It

follows that $a_{12}\overline{a}_{12}+\cdots+a_{1n}\overline{a}_{1n} = 0$. Since $a_{1j}\overline{a}_{1j} \geq 0$, this shows that the entries a_{1j} with $j > 1$ are zero and that

$$M_1 = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & N & \\ 0 & & & \end{bmatrix}.$$

We continue, working on N. □

8. SKEW-SYMMETRIC FORMS

The theory of skew-symmetric forms is independent of the field of scalars. One might expect trouble with fields of characteristic 2, in which $1 + 1 = 0$. They look peculiar because $a = -a$ for all a, so the conditions for symmetry (1.5) and for skew symmetry (1.6) are the same. It turns out that fields of characteristic 2 don't cause trouble with skew-symmetric forms, if the definition of skew symmetry is changed to handle them. The definition which works for all fields is this:

(8.1) **Definition.** A bilinear form $\langle\,,\rangle$ on a vector space V is *skew-symmetric* if

$$\langle v, v\rangle = 0$$

for all $v \in V$.

The rule

(8.2) $$\langle v, w\rangle = -\langle w, v\rangle$$

for all $v, w \in V$ continues to hold with this definition. It is proved by expanding

$$\langle v + w, v + w\rangle = \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle,$$

and by using the fact that $\langle v, v\rangle = \langle w, w\rangle = \langle v + w, v + w\rangle = 0$. If the characteristic of the field of scalars is not 2, then (8.1) and (8.2) are equivalent. For if (8.2) holds for all v, w, then setting $w = v$ we find $\langle v, v\rangle = -\langle v, v\rangle$. This implies that $2\langle v, v\rangle = 0$, hence that $\langle v, v\rangle = 0$ unless $2 = 0$ in the field.

Note that if F has characteristic 2, then $1 = -1$ in F, so (8.2) shows that the form is actually symmetric. But most symmetric forms don't satisfy (8.1).

The matrix A of a skew-symmetric form with respect to an arbitrary basis is characterized by the properties

(8.3) $$a_{ii} = 0 \quad \text{and} \quad a_{ij} = -a_{ji}, \quad \text{if } i \neq j.$$

We take these properties as the definition of a *skew-symmetric matrix*. If the characteristic is not 2, then this is equivalent with the condition

(8.4) $$A^t = -A.$$

(8.5) **Theorem.**

(a) Let V be a vector space of dimension m over a field F, and let $\langle\, ,\rangle$ be a nondegenerate skew-symmetric form on V. Then m is an even integer, and there is a basis **B** of V such that the matrix A of the form with respect to that basis is

$$J_{2n} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix},$$

where $0, I$ denote the $n \times n$ matrices and $n = \frac{1}{2}m$.

(b) *Matrix form:* Let A be a nonsingular skew-symmetric $m \times m$ matrix. Then m is even, and there is a matrix $Q \in GL_m(F)$ such that QAQ^t is the matrix J_{2n}.

A basis **B** as in (8.6a) is called a *standard symplectic basis*. Note that rearranging the standard symplectic basis in the order $(v_1, v_{n+1}, v_2, v_{n+2}, \dots, v_n, v_{2n})$ changes the matrix J_{2n} into a matrix made up of 2×2 blocks

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

along the diagonal. This is the form which is most convenient for proving the theorem. We leave the proof as an exercise. □

9. SUMMARY OF RESULTS, IN MATRIX NOTATION

Real numbers: A square matrix A is *symmetric* if $A^t = A$ and *orthogonal* if $A^t = A^{-1}$.

(1) *Spectral Theorem:* If A is a real symmetric matrix, there is an orthogonal matrix P such that $PAP^t (= PAP^{-1})$ is diagonal.

(2) If A is a real symmetric matrix, there is a real invertible matrix P such that

$$PAP^t = \begin{bmatrix} I_p & & \\ & -I_m & \\ & & 0_z \end{bmatrix},$$

for some integers p, m, z.

(3) *Sylvester's Law:* The numbers p, m, z are determined by the matrix A.

Complex numbers: A complex square matrix A is *hermitian* if $A^* = A$, *unitary* if $A^* = A^{-1}$, and *normal* if $AA^* = A^*A$.

(1) *Spectral Theorem*: If A is a hermitian matrix, there is a unitary matrix P such that PAP^* is a real diagonal matrix.

(2) If A is a normal matrix, there is a unitary matrix P such that PAP^* is diagonal.

F arbitrary: A square $n \times n$ matrix is *skew-symmetric* if $a_{ii} = 0$ and $a_{ij} = -a_{ji}$ for all i, j. If A is an invertible skew-symmetric matrix, then n is even, and there is an invertible matrix P so that PAP^t has the form

$$\begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}.$$

(9.1) *Note.* The rule $A' = (P^t)^{-1}A(P^{-1})$ for change of basis in a bilinear form (see (1.12)) is rather ugly because of the way the matrix P of change of coordinates is defined. It is possible to rearrange equations (4.17) of Chapter 3, by writing

$$(9.2) \qquad v_i' = \sum_j q_{ij}v_j \quad \text{or} \quad \mathbf{B}'^t = Q\mathbf{B}^t.$$

This results in $Q = (P^{-1})^t$, and with this rule we obtain the nicer formula

$$A' = QAQ^t,$$

to replace (1.12). We can use it if we want to.

The problem with formula (9.2) is that change of basis on a linear transformation gets messed up; namely the formula $A' = PAP^{-1}$ [Chapter 4 (3.4)] is replaced by $A' = (Q^{-1})^t AQ^t$. Trying to keep the formulas neat is like trying to smooth a bump in a rug.

This brings up an important point. Linear operators on V and bilinear forms on V are each given by an $n \times n$ matrix A, once a basis has been chosen. One is tempted to think that the theories of linear operators and of bilinear forms are somehow equivalent, but they are not, unless a basis is fixed. For under a *change* of basis the matrix of a bilinear form changes to $(P^t)^{-1}AP^{-1}$ (1.12), while the matrix of a linear operator changes to PAP^{-1} [Chapter 4 (3.4)]. So the new matrices are no longer equal. To be precise, this shows that the theories diverge when the basis is changed, unless the matrix P of change of basis happens to be orthogonal. If P is orthogonal, then $P = (P^t)^{-1}$, and we are all right. The matrices remain equal. This is one benefit of working with orthonormal bases.



Yvonne Verdier

EXERCISES

1. Definition of Bilinear Form

1. Let A and B be real $n \times n$ matrices. Prove that if $X^tAY = X^tBY$ for all vectors X, Y in $\mathbb{R}^n$, then $A = B$.

2. Prove directly that the bilinear form represented by the matrix $\begin{bmatrix} a & b \\ b & d \end{bmatrix}$ is positive definite if and only if $a > 0$ and $ad - b^2 > 0$.

3. Apply the Gram–Schmidt procedure to the basis $(1,1,0)^t$, $(1,0,1)^t$, $(0,1,1)^t$, when the form is dot product.

4. Let $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. Find an orthonormal basis for $\mathbb{R}^2$ with respect to the form X^tAY.

5. **(a)** Prove that every real square matrix is the sum of a symmetric matrix and a skew-symmetric matrix ($A^t = -A$) in exactly one way.

(b) Let $\langle\, , \rangle$ be a bilinear form on a real vector space V. Show that there is a symmetric form $(\, ,)$ and a skew-symmetric form $[\, ,]$ so that $\langle\, , \rangle = (\, ,) + [\, ,]$.

6. Let $\langle\, , \rangle$ be a symmetric bilinear form on a vector space V over a field F. The function q: $V \longrightarrow F$ defined by $q(v) = \langle v, v \rangle$ is called the *quadratic form* associated to the bilinear form. Show how to recover the bilinear form from q, if the characteristic of the field F is not 2, by expanding $q(v + w)$.

***7.** Let X, Y be vectors in $\mathbb{C}^n$, and assume that $X \neq 0$. Prove that there is a symmetric matrix B such that $BX = Y$.

2. *Symmetric Forms: Orthogonality*

1. Prove that a positive definite form is nondegenerate.

2. A matrix A is called *positive semidefinite* if $X^tAX \geq 0$ for all $X \in \mathbb{R}^n$. Prove that A^tA is positive semidefinite for any $m \times n$ real matrix A.

3. Find an orthogonal basis for the form on $\mathbb{R}^n$ whose matrix is as follows.

(a) $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

4. Extend the vector $X_1 = (1,1,1)^t/\sqrt{3}$ to an orthonormal basis for $\mathbb{R}^3$.

***5.** Prove that if the columns of an $n \times n$ matrix A form an orthonormal basis, then the rows do too.

6. Let A, A' be symmetric matrices related by $A = P^tA'P$, where $P \in GL_n(F)$. Is it true that the ranks of A and of A' are equal?

7. Let A be the matrix of a symmetric bilinear form $\langle\, , \rangle$ with respect to some basis. Prove or disprove: The eigenvalues of A are independent of the basis.

8. Prove that the only real matrix which is orthogonal, symmetric, and positive definite is the identity.

9. The vector space P of all real polynomials of degree $\leq n$ has a bilinear form, defined by

$$\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)dx.$$

Find an orthonormal basis for P when n has the following values. **(a)** 1 **(b)** 2 **(c)** 3

10. Let V denote the vector space of real $n \times n$ matrices. Prove that $\langle A, B \rangle = \operatorname{trace}(A^tB)$ is a positive definite bilinear form on V. Find an orthonormal basis for this form.

11. A symmetric matrix A is called *negative definite* if $X^tAX < 0$ for all $X \neq 0$. Give a criterion analogous to (1.26) for a symmetric matrix A to be negative definite.

12. Prove that every symmetric nonsingular complex matrix A has the form $A = P^tP$.

13. In the notation of (2.12), show by example that the span of $(v_1, \ldots, v_p)$ is not determined by the form.

14. **(a)** Let W be a subspace of a vector space V on which a symmetric bilinear form is given. Prove that $W^\perp$ is a subspace.
(b) Prove that the null space N is a subspace.

15. Let W_1, W_2 be subspaces of a vector space V with a symmetric bilinear form. Prove each of the following.
(a) $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$ **(b)** $W \subset W^{\perp\perp}$ **(c)** If $W_1 \subset W_2$, then $W_1^\perp \supset W_2^\perp$.

16. Prove Proposition (2.7), that $V = W \oplus W^\perp$ if the form is nondegenerate on W.

17. Let $V = \mathbb{R}^{2\times 2}$ be the vector space of real 2×2 matrices.
(a) Determine the matrix of the bilinear form $\langle A, B\rangle = \text{trace}(AB)$ on V with respect to the standard basis $\{e_{ij}\}$.
(b) Determine the signature of this form.
(c) Find an orthogonal basis for this form.
(d) Determine the signature of the form on the subspace of V of matrices with trace zero.

***18.** Determine the signature of the form $\langle A, B\rangle = \text{trace } AB$ on the space $\mathbb{R}^{n\times n}$ of real $n \times n$ matrices.

19. Let $V = \mathbb{R}^{2\times 2}$ be the space of 2×2 matrices.
(a) Show that the form $\langle A, B\rangle$ defined by $\langle A, B\rangle = \det(A + B) - \det A - \det B$ is symmetric and bilinear.
(b) Compute the matrix of this form with respect to the standard basis $\{e_{ij}\}$, and determine the signature of the form.
(c) Do the same for the subspace of matrices of trace zero.

20. Do exercise 19 for $\mathbb{R}^{3\times 3}$, replacing the quadratic form $\det A$ by the coefficient of t in the characteristic polynomial of A.

21. Decide what the analogue of Sylvester's Law for symmetric forms over complex vector spaces is, and prove it.

22. Using the method of proof of Theorem (2.9), find necessary and sufficient conditions on a field F so that every finite-dimensional vector space V over F with a symmetric bilinear form $\langle\, ,\rangle$ has an orthogonal basis.

23. Let $F = \mathbb{F}_2$, and let $A = \begin{bmatrix} & 1 \\ 1 & \end{bmatrix}$.
(a) Prove that the bilinear form X^tAY on F^2 can not be diagonalized.
(b) Determine the orbits for the action $P, A \rightsquigarrow PAP^t$ of $GL_2(F)$ on the space of 2×2 matrices with coefficients in F.

3. *The Geometry Associated to a Positive Form*

1. Let V be a Euclidean space. Prove the Schwarz Inequality and the Triangle Inequality.

2. Let W be a subspace of a Euclidean space V. Prove that $W = W^{\perp\perp}$.

3. Let V be a Euclidean space. Show that if $|v| = |w|$, then $(v + w) \perp (v - w)$. Interpret this formula geometrically.

4. Prove the parallelogram law $|v + w|^2 + |v - w|^2 = 2|v|^2 + 2|w|^2$ in a Euclidean space.
5. Prove that the orthogonal projection (3.7) is a linear transformation.
6. Find the matrix of the projection $\pi: \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ such that the image of the standard bases of $\mathbb{R}^3$ forms an equilateral triangle and $\pi(e_1)$ points in the direction of the x-axis.

*7. Let W be a two-dimensional subspace of $\mathbb{R}^3$, and consider the orthogonal projection π of $\mathbb{R}^3$ onto W. Let $(a_i, b_i)^t$ be the coordinate vector of $\pi(e_i)$, with respect to a chosen orthonormal basis of W. Prove that (a_1, a_2, a_3) and (b_1, b_2, b_3) are orthogonal unit vectors.

*8. Let $w \in \mathbb{R}^n$ be a vector of length 1.
   (a) Prove that the matrix $P = I - 2ww^t$ is orthogonal.
   (b) Prove that multiplication by P is a reflection through the space W orthogonal to w, that is, prove that if we write an arbitrary vector v in the form $v = cw + w'$, where $w' \in W^\perp$, then $Pv = -cw + w'$.
   (c) Let X, Y be arbitrary vectors in $\mathbb{R}^n$ with the same length. Determine a vector w such that $PX = Y$.

*9. Use exercise 8 to prove that every orthogonal $n \times n$ matrix is a product of at most n reflections.

10. Let A be a real symmetric matrix, and let T be the linear operator on $\mathbb{R}^n$ whose matrix is A.
    (a) Prove that $(\ker T) \perp (\operatorname{im} T)$ and that $V = (\ker T) \oplus (\operatorname{im} T)$.
    (b) Prove that T is an orthogonal projection onto im T if and only if, in addition to being symmetric, $A^2 = A$.
11. Let A be symmetric and positive definite. Prove that the maximal matrix entries are on the diagonal.

4. *Hermitian Forms*

1. Verify rules (4.4).
2. Show that the dot product form $(X \cdot Y) = X^t Y$ is not positive definite on $\mathbb{C}^n$.
3. Prove that a matrix A is hermitian if and only if the associated form X^*AX is a hermitian form.
4. Prove that if X^*AX is real for all complex vectors X, then A is hermitian.
5. Prove that the $n \times n$ hermitian matrices form a real vector space, and find a basis for that space.
6. Let V be a two-dimensional hermitian space. Let (v_1, v_2) be an orthonormal basis for V. Describe all orthonormal bases (v_1', v_2') with $v_1 = v_1'$.
7. Let $X, Y \in \mathbb{C}^n$ be orthogonal vectors. Prove that $|X + Y|^2 = |X|^2 + |Y|^2$.
8. Is $\langle X, Y \rangle = x_1 y_1 + i x_1 y_2 - i x_2 y_1 + i x_2 y_2$ on $\mathbb{C}^2$ a hermitian form?
9. Let A, B be positive definite hermitian matrices. Determine which of the following matrices are positive definite hermitian: A^2, A^{-1}, AB, $A + B$.
10. Prove that the determinant of a hermitian matrix is a real number.
11. Prove that A is positive definite hermitian if and only if $A = P^*P$ for some invertible matrix P.
12. Prove Theorem (4.19), that a hermitian form on a complex vector space V has an orthonormal basis if and only if it is positive definite.

13. Extend the criterion (1.26) for positive definiteness to hermitian matrices.

14. State and prove an analogue of Sylvester's Law for hermitian matrices.

15. Let $\langle\,,\rangle$ be a hermitian form on a complex vector space V, and let $\{v, w\}$ denote the real part of the complex number $\langle v, w\rangle$. Prove that if V is regarded as a real vector space, then $\{\,,\}$ is a symmetric bilinear form on V, and if $\langle\,,\rangle$ is positive definite, then $\{\,,\}$ is too. What can you say about the imaginary part?

16. Let P be the vector space of polynomials of degree $\leq n$.

(a) Show that

$$\langle f, g\rangle = \int_0^{2\pi} \overline{f(e^{i\theta})}g(e^{i\theta})d\theta$$

is a positive definite hermitian form on P.

(b) Find an orthonormal basis for this form.

17. Determine whether or not the following rules define hermitian forms on the space $\mathbb{C}^{n\times n}$ of complex matrices, and if so, determine their signature.

(a) $A, B \rightsquigarrow \text{trace}(A^*B)$ (b) $A, B \rightsquigarrow \text{trace}(\overline{A}B)$

18. Let A be a unitary matrix. Prove that $|\det A| = 1$.

19. Let P be a unitary matrix, and let X_1, X_2 be eigenvectors for P, with distinct eigenvalues λ_1, λ_2. Prove that X_1 and X_2 are orthogonal with respect to the standard hermitian product on $\mathbb{C}^n$.

***20.** Let A be any complex matrix. Prove that $I + A^*A$ is nonsingular.

21. Prove Proposition (4.20).

5. *The Spectral Theorem*

1. Prove that if T is a hermitian operator then the rule $\{v, w\} = \langle v, Tw\rangle = X^*MY$ defines a second hermitian form on V.

2. Prove that the eigenvalues of a real symmetric matrix are real numbers.

3. Prove that eigenvectors associated to distinct eigenvalues of a hermitian matrix A are orthogonal.

4. Find a unitary matrix P so that PAP^* is diagonal, when

$$A = \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}.$$

5. Find a real orthogonal matrix P so that PAP^t is diagonal, when

$$\textbf{(a)}\ A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad \textbf{(b)}\ A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad \textbf{(c)}\ A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

6. Prove the equivalence of conditions (a) and (b) of the Spectral Theorem.

7. Prove that a real symmetric matrix A is positive definite if and only if its eigenvalues are positive.

8. Show that the only matrix which is both positive definite hermitian and unitary is the identity I.

9. Let A be a real symmetric matrix. Prove that e^A is symmetric and positive definite.

10. Prove that for any square matrix A, $\ker A = (\operatorname{im} A^*)^{\perp}$.

***11.** Let $\zeta = e^{2\pi i/n}$, and let A be the $n \times n$ matrix $a_{jk} = \zeta^{jk}/\sqrt{n}$. Prove that A is unitary.

12. Show that for every complex matrix A there is a unitary matrix P such that PAP^* is upper triangular.

13. Let A be a hermitian matrix. Prove that there is a unitary matrix P with determinant 1 such that PAP^* is diagonal.

***14.** Let A, B be hermitian matrices which commute. Prove that there is a unitary matrix P such that PAP^* and PBP^* are both diagonal.

15. Use the Spectral Theorem to give a new proof of the fact that a positive definite real symmetric $n \times n$ matrix P has the form $P = AA^t$ for some $n \times n$ matrix A.

16. Let λ, μ be distinct eigenvalues of a complex symmetric matrix A, and let X, Y be eigenvectors associated to these eigenvalues. Prove that X is orthogonal to Y with respect to dot product.

6. Conics and Quadrics

1. Determine the type of the quadric $x^2+4xy+2xz+z^2 + 3x+z - 6 = 0$.

2. Suppose that (6.1) represents an ellipse. Instead of diagonalizing the form and then making a translation to reduce to the standard type, we could make the translation first. Show how to compute the required translation by calculus.

3. Discuss all degenerate loci for conics.

4. Give a necessary and sufficient condition, in terms of the coefficients of its equation, for a conic to be a circle.

5. (a) Describe the types of conic in terms of the signature of the quadratic form.
(b) Do the same for quadrics in $\mathbb{R}^3$.

6. Describe the degenerate quadrics, that is, those which are not listed in (6.14).

7. The Spectral Theorem for Normal Operators

1. Show that for any normal matrix A, $\ker A = (\operatorname{im} A)^{\perp}$.

2. Prove or disprove: If A is a normal matrix and W is an A-invariant subspace of $V = \mathbb{C}^n$, then $W^{\perp}$ is also A-invariant.

3. A matrix is skew-hermitian if $A^* = -A$. What can you say about the eigenvalues and the possibility of diagonalizing such a matrix?

4. Prove that the cyclic shift operator

$$\begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & 1 \\ 1 & & & 0 \end{bmatrix}$$

is normal, and determine its diagonalization.

5. Let P be a real matrix which is normal and has real eigenvalues. Prove that P is symmetric.

6. Let P be a real skew-symmetric matrix. Prove that P is normal.

*7. Prove that the *circulant*

$$\begin{bmatrix} c_0 & c_1 & c_2 & \cdots & c_n \\ c_n & c_0 & c_1 & \cdots & c_{n-1} \\ \vdots & & & & \vdots \\ c_1 & c_2 & c_3 & \cdots & c_0 \end{bmatrix}$$

is a normal matrix.

8. (a) Let A be a complex symmetric matrix. Prove that eigenvectors of A with distinct eigenvalues are orthogonal with respect to the bilinear form X^tX.
   *(b) Give an example of a complex symmetric matrix A such that there is no $P \in O_n(\mathbb{C})$ with PAP^t diagonal.
9. Let A be a normal matrix. Prove that A is hermitian if and only if all eigenvalues of A are real, and that A is unitary if and only if every eigenvalue has absolute value 1.
10. Let V be a finite-dimensional complex vector space with a positive definite hermitian form $\langle\, ,\rangle$, and let $T: V \longrightarrow V$ be a linear operator on V. Let A be the matrix of T with respect to an orthonormal basis $\mathbf{B}$. The adjoint operator $T^*: V \longrightarrow V$ is defined as the operator whose matrix with respect to the same basis is A^*.
    (a) Prove that T and T^* are related by the equations $\langle Tv, w\rangle = \langle v, T^*w\rangle$ and $\langle v, Tw\rangle = \langle T^*v, w\rangle$ for all $v, w \in W$. Prove that the first of these equations characterizes T^*.
    (b) Prove that T^* does not depend on the choice of orthonormal basis.
    (c) Let v be an eigenvector for T with eigenvalue λ, and let $W = v^\perp$ be the space of vectors orthogonal to v. Prove that W is T^*-invariant.
11. Prove that for any linear operator T, TT^* is hermitian.
12. Let V be a finite-dimensional complex vector space with a positive definite hermitian form $\langle\, ,\rangle$. A linear operator $T: V \longrightarrow V$ is called *normal* if $TT^* = T^*T$.
    (a) Prove that T is normal if and only if $\langle Tv, Tw\rangle = \langle T^*v, T^*w\rangle$ for all $v, w \in V$, and verify that hermitian operators and unitary operators are normal.
    (b) Assume that T is a normal operator, and let v be an eigenvector for T, with eigenvalue λ. Prove that v is also an eigenvector for T^*, and determine its eigenvalue.
    (c) Prove that if v is an eigenvector, then $W = v^\perp$ is T-invariant, and use this to prove the Spectral Theorem for normal operators.

8. Skew-Symmetric Forms

1. Prove or disprove: A matrix A is skew-symmetric if and only if $X^tAX = 0$ for all X.
2. Prove that a form is skew-symmetric if and only if its matrix has the properties (8.4).
3. Prove or disprove: A skew-symmetric $n \times n$ matrix is singular if n is odd.
4. Prove or disprove: The eigenvalues of a real skew-symmetric matrix are purely imaginary.

*5. Let S be a real skew-symmetric matrix. Prove that $I + S$ is invertible, and that $(I - S)(I + S)^{-1}$ is orthogonal.

*6. Let A be a real skew-symmetric matrix.
   (a) Prove that $\det A \geq 0$.
   (b) Prove that if A has integer entries, then $\det A$ is the square of an integer.

7. Let $\langle\, ,\rangle$ be a skew-symmetric form on a vector space V. Define orthogonality, null space, and nondegenerate forms as in Section 2.
   (a) Prove that the form is nondegenerate if and only if its matrix with respect to any basis is nonsingular.
   (b) Prove that if W is a subspace such that the restriction of the form to W is nondegenerate, then $V = W \oplus W^{\perp}$.
   (c) Prove that if the form is not identically zero, then there is a subspace W and a basis of W such that the restriction of the form to W has matrix $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.
   (d) Prove Theorem (8.6).

9. Summary of Results, in Matrix Notation

1. Determine the symmetry of the matrices $AB + BA$ and $AB - BA$ in the following cases.
   (a) A,B symmetric **(b)** A,B hermitian **(c)** A,B skew-symmetric **(d)** A symmetric, B skew-symmetric
2. State which of the following rules define operations of $GL_n(\mathbb{C})$ on the space $\mathbb{C}^{n\times n}$ of all complex matrices:

$$P, A \rightsquigarrow PAP^t,\ (P^{-1})^t A(P^{-1}),\ (P^{-1})^t AP^t,\ P^{-1}AP^t,\ AP^t,\ P^tA.$$

3. **(a)** With each of the following types of matrices, describe the possible determinants:
   (i) real orthogonal **(ii)** complex orthogonal **(iii)** unitary **(iv)** hermitian
   (v) symplectic **(vi)** real symmetric, positive definite **(vii)** real symmetric, negative definite
   (b) Which of these types of matrices have only real eigenvalues?
4. **(a)** Let E be an arbitrary complex matrix. Prove that the matrix $\begin{bmatrix} I & E^* \\ -E & I \end{bmatrix}$ is invertible.
   (b) Find the inverse in block form $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$.

*5. **(a)** What is wrong with the following argument? Let P be a real orthogonal matrix. Let X be a (possibly complex) eigenvector of P, with eigenvalue λ. Then $X^tP^tX = (PX)^tX = \lambda X^tX$. On the other hand, $X^tP^tX = X^t(P^{-1}X) = \lambda^{-1}X^tX$. Therefore $\lambda = \lambda^{-1}$, and so $\lambda = \pm 1$.
   (b) State and prove a correct theorem based on this argument.

*6. Show how to describe any element of SO_4 in terms of rotations of two orthogonal planes in $\mathbb{R}^4$.

*7. Let A be a real $m \times n$ matrix. Prove that there are orthogonal matrices $P \in O_m$ and $Q \in O_n$ such that $PAQ = D$ is diagonal, with nonnegative diagonal entries.

Chapter 8

Linear Groups

In these days the angel of topology and the devil of abstract algebra fight for the soul of every individual discipline of mathematics.

Hermann Weyl

1. THE CLASSICAL LINEAR GROUPS

Subgroups of the general linear group GL_n are called *linear groups*. In this chapter we will study the most important ones: the orthogonal, unitary, and symplectic groups. They are called the *classical groups*.

The classical groups arise as stabilizers for some natural operations of GL_n on the space of $n \times n$ matrices. The first of these operations is that which describes change of basis in a bilinear form. The rule

$$(1.1) \qquad P, A \rightsquigarrow (P^t)^{-1}AP^{-1}$$

is an operation of GL_n on the set of all $n \times n$ matrices. This is true for any field of scalars, but we will be interested in the real and complex cases. As we have seen in Chapter 7 (1.15), the orbit of a matrix A under this operation is the set of matrices A' which represent the form X^tAY, but with respect to different bases. It is customary to call matrices in the same orbit *congruent*. We can set $Q = (P^t)^{-1}$ to obtain the equivalent definition

(1.2) *A and A' are congruent if $A' = QAQ^t$ for some $Q \in GL_m(F)$.*

Sylvester's Law [Chapter 7 (2.11)] describes the different orbits or congruence classes of real symmetric matrices. Every congruence class of real symmetric matrices contains exactly one matrix of the form Chapter 7 (2.10). The *orthogonal group,* which we have defined before, is the stabilizer of the identity matrix for this operation. As before, we will denote the real orthogonal group by the symbol O_n:

$$(1.3) \qquad O_n = \{P \in GL_n(\mathbb{R}) \mid P^tP = I\}.$$

The complex orthogonal group is defined analogously:

$$O_n(\mathbb{C}) = \{P \in GL_n(\mathbb{C}) \mid P^t P = I\}.$$

The stabilizer of the *Lorentz form* [Chapter 7 (2.16)], defined by the matrix

$$I_{3,1} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{bmatrix},$$

is called the *Lorentz group*. It is denoted by $O_{3,1}(\mathbb{R})$ or $O_{3,1}$:

(1.4) $$O_{3,1} = \{P \in GL_n(\mathbb{R}) \mid P^t I_{3,1} P = I_{3,1}\}.$$

The linear operators represented by these matrices are often called *Lorentz transformations*. The subscript (3, 1) indicates the signature of the matrix, the number of $+1$'s and -1's. In this way an analogous group $O_{p,q}$ can be defined for any signature (p, q).

The operation (1.1) also describes change of basis in forms which are not symmetric. Thus Theorem (8.6) of Chapter 7 tells us this:

(1.5) **Corollary.** There is exactly one congruence class of real nonsingular skew-symmetric $m \times m$ matrices, if m is even. □

The standard skew-symmetric form is defined by the $2n \times 2n$ matrix J (Chapter 7 (8.5)), and its stabilizer is called the *symplectic group*

(1.6) $$SP_{2n}(\mathbb{R}) = \{P \in GL_{2n}(\mathbb{R}) \mid P^t J P = J\}.$$

Again, the complex symplectic group $SP_{2n}(\mathbb{C})$ is defined analogously.

Finally, the *unitary group* is defined in terms of the operation

(1.7) $$P, A \rightsquigarrow (P^*)^{-1} A P^{-1}.$$

This definition makes sense only when the field of scalars is the complex field. Exactly as with bilinear forms, the orbit of a matrix A consists of the matrices which define the form $\langle X, Y\rangle = X^*AY$ with respect to different bases (see [Chapter 7 (4.12)]). The unitary group is the stabilizer of the identity matrix for this action:

(1.8) $$U_n = \{P \mid P^*P = I\}.$$

Thus U_n is the group of matrices representing changes of basis which leaves the hermitian dot product [Chapter 7 (4.2)] X^*Y invariant.

The word *special* is added to indicate the subgroup of matrices with determinant 1. This gives us some more groups:

Special linear group	$SL_n(\mathbb{R})$: $n \times n$ matrices P with determinant 1;
Special orthogonal group	$SO_n(\mathbb{R})$: the intersection $SL_n(\mathbb{R}) \cap O_n(\mathbb{R})$;
Special unitary group	SU_n: the intersection $SL_n(\mathbb{C}) \cap U_n$.

Though this is not obvious from the definition, symplectic matrices have determinant 1, so the two uses of the letter S do not cause conflict.

2. THE SPECIAL UNITARY GROUP SU_2

The main object of this chapter is to describe the geometric properties of the classical linear groups, by considering them as subsets of the spaces $\mathbb{R}^{n\times n}$ or $\mathbb{C}^{n\times n}$ of all matrices. We know the geometry of a few groups already. For example, $GL_1(\mathbb{C}) = \mathbb{C}^\times$ is the "punctured plane" $\mathbb{C} - \{0\}$. Also, if p is a 1×1 matrix, then $p^* = \overline{p}$. Thus

$$U_1 = \{p \in \mathbb{C}^\times \mid \overline{p}p = 1\}. \tag{2.1}$$

This is the set of complex numbers of absolute value 1—the unit circle in the complex plane. We can identify it with the unit circle in $\mathbb{R}^2$,

$$x_1^2 + x_2^2 = 1,$$

by sending $x_1 + x_2 i \rightsquigarrow (x_1, x_2)$. The group SO_2 of rotations of the plane is isomorphic to U_1. It is also a circle, embedded into $\mathbb{R}^{2\times 2}$ by the map

$$(x_1, x_2) \rightsquigarrow \begin{bmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{bmatrix}. \tag{2.2}$$

We will describe some more of the groups in the following sections.

The *dimension* of a linear group G is, roughly speaking, the number of degrees of freedom of a matrix in G. The group SO_2, for example, has dimension 1. A matrix in SO_2 represents rotation by an angle θ, and this angle is the single parameter needed to determine the rotation. We will discuss dimension more carefully in Section 7, but we want to describe some of the low-dimensional groups explicitly first. The smallest dimension in which really interesting groups appear is 3, and three of these—SU_2, SO_3, and $SL_2(\mathbb{R})$—are very important. We will study the special unitary group SU_2 in this section.

Let $P = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be an element of SU_2, with $a, b, c, d \in \mathbb{C}$. The equations defining SU_2 are $P^*P = I$ and $\det P = 1$. By Cramer's Rule,

$$P^{-1} = (\det P)^{-1} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Since $P^{-1} = P^*$ for a matrix in SU_2, we find $\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} \overline{a} & \overline{c} \\ \overline{b} & \overline{d} \end{bmatrix}$, or

$$\overline{a} = d, \quad \text{and} \quad \overline{b} = -c. \tag{2.3}$$

Thus

$$P = \begin{bmatrix} a & b \\ -\overline{b} & \overline{a} \end{bmatrix}. \tag{2.4}$$

The condition $\det P = 1$ has become lost in the computation and must be put back:

$$\bar{a}a + \bar{b}b = 1. \tag{2.5}$$

Equations (2.3) and (2.5) provide a complete list of conditions describing the entries of a matrix in SU_2. The matrix P is described by the vector $(a, b) \in \mathbb{C}^2$ of length 1, and any such vector gives us a matrix $P \in SU_2$ by the rule (2.4).

If we write out a, b in terms of their real and imaginary parts, equation (2.5) gives us a bijective correspondence between SU_2 and points of $\mathbb{R}^4$ lying on the locus

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1. \tag{2.6}$$

This equation is equivalent to (2.5) if we set $a = x_1 + x_2 i$ and $b = x_3 + x_4 i$.

The locus (2.6) is called the *unit* 3-*sphere* in $\mathbb{R}^4$, in analogy with the unit sphere in $\mathbb{R}^3$. The number 3 refers to its dimension, the number of degrees of freedom of a point on the sphere. Thus the unit sphere

$$x_1^2 + x_2^2 + x_3^2 = 1$$

in $\mathbb{R}^3$, being a surface, is called a *2-sphere*. The unit circle in $\mathbb{R}^2$, a curve, is called a *1-sphere*. We will sometimes denote a sphere of dimension d by S^d.

A bijective map $f: S \longrightarrow S'$ between subsets of Euclidean spaces is called a *homeomorphism* if f and f^{-1} are continuous maps (Appendix, Section 3). The correspondence between SU_2, considered as a subset of $\mathbb{C}^{2\times 2}$, and the sphere (2.6) is obviously continuous, as is its inverse. Therefore these two spaces are homeomorphic.

(2.7) *SU_2 is homeomorphic to the unit* 3-*sphere in* $\mathbb{R}^4$.

It is convenient to identify SU_2 with the 3-sphere. We can do this if we represent the matrix (2.4) by its top row, the vector $(a, b) \in \mathbb{C}^2$, or by the vector $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$. These representations can be thought of as different notations for the same element P of the group, and we will pass informally from one representation to the other. For geometric visualization, the representations $P = (a, b)$ and $P = (x_1, x_2, x_3, x_4)$, being in lower-dimensional spaces, are more convenient.

The fact that the 3-sphere has a group structure is remarkable, because there is no way to make the 2-sphere into a group with a continuous law of composition. In fact, a famous theorem of topology asserts that the only spheres with continuous group laws are the 1-sphere, which is realized as the rotation group SO_2, and the 3-sphere SU_2.

We will now describe the algebraic structures on SU_2 analogous to the curves of constant latitude and longitude on the 2-sphere. The matrices $I, -I$ will play the roles of the north and south poles. In our vector notation, they are the points $(\pm 1, 0, 0, 0)$ of the sphere.

If the poles of the 2-sphere $x_1^2 + x_2^2 + x_3^2 = 1$ are placed at the points $(\pm 1, 0, 0)$, then the latitudes are the circles $x_1 = c$, $-1 < c < 1$. The analogues on the 3-sphere SU_2 of these latitudes are the surfaces on which the x_1-coordinate is constant. They are two-dimensional spheres, embedded into $\mathbb{R}^4$ by

$$x_1 = c \quad \text{and} \quad x_2^2 + x_3^2 + x_4^2 = (1 - c^2), \quad -1 < c < 1. \tag{2.8}$$

These sets can be described algebraically as *conjugacy classes* in SU_2.

(2.9) **Proposition.** Except for two special classes, the conjugacy classes in SU_2 are the *latitudes*, the sets defined by the equations (2.8). For a given c in the interval $(-1, 1)$, this set consists of all matrices $P \in SU_2$ such that trace $P = 2c$. The remaining conjugacy classes are $\{I\}$ and $\{-I\}$, each consisting of one element. These two classes make up the center $Z = \{\pm I\}$ of the group SU_2.

Proof. The characteristic polynomial of the matrix P (2.4) is

$$t^2 - (a + \overline{a})t + 1 = t^2 - 2x_1 t + 1. \tag{2.10}$$

This polynomial has a pair $\lambda, \overline{\lambda}$ of complex conjugate roots on the unit circle, and the roots, the eigenvalues of P, depend only on trace $P = 2x_1$. Furthermore, two matrices with different traces have different eigenvalues. The proposition will follow if we show that the conjugacy class of P contains every matrix in SU_2 with the same eigenvalues. The cases $x_1 = 1, -1$ correspond to the two special conjugacy classes $\{I\}, \{-I\}$, so the proof is completed by the next lemma.

(2.11) **Lemma.** Let P be an element of SU_2, with eigenvalues $\lambda, \overline{\lambda}$. Then P is conjugate in SU_2 to the matrix

$$\begin{bmatrix} \lambda & \\ & \overline{\lambda} \end{bmatrix}.$$

Proof. By the Spectral Theorem for normal operators [Chapter 7 (7.3)], there is a unitary matrix Q so that QPQ^* is diagonal. We only have to show that Q can be chosen so as to have determinant 1. Say that $\det Q = \delta$. Since $Q^*Q = I$, $(\det Q^*)(\det Q) = \overline{\delta}\delta = 1$; hence δ has absolute value 1. Let ϵ be a square root of δ. Then $\overline{\epsilon}\epsilon = 1$ too. The matrix $Q_1 = \overline{\epsilon}Q$ is in SU_2, and $P_1 = Q_1 P Q_1^*$ is also diagonal. The diagonal entries of P_1 are the eigenvalues $\lambda, \overline{\lambda}$. The eigenvalues can be interchanged, if desired, by conjugating by the matrix

$$Q_2 = \begin{bmatrix} & 1 \\ -1 & \end{bmatrix}, \tag{2.12}$$

which is also an element of SU_2. □

Next we will introduce the longitudes of SU_2. The longitudes on the 2-sphere $x_1^2 + x_2^2 + x_3^2 = 1$ can be described as intersections of the sphere with planes containing the two poles $(\pm 1, 0, 0)$. When we add a fourth variable x_4 to get the equation of the 3-sphere, a natural way to extend this definition is to form the intersection with a two-dimensional subspace of $\mathbb{R}^4$ containing the two poles $\pm I$. This is a circle in SU_2, and we will think of these circles as the longitudes. Thus while the latitudes on SU_2 are 2-spheres, the *longitudes* are 1-spheres, the "great circles" through the poles.

Note that every point $P = (x_1, x_2, x_3, x_4)$ of SU_2 except for the poles is contained in exactly one longitude. This is because if P is not a pole, then P and I will

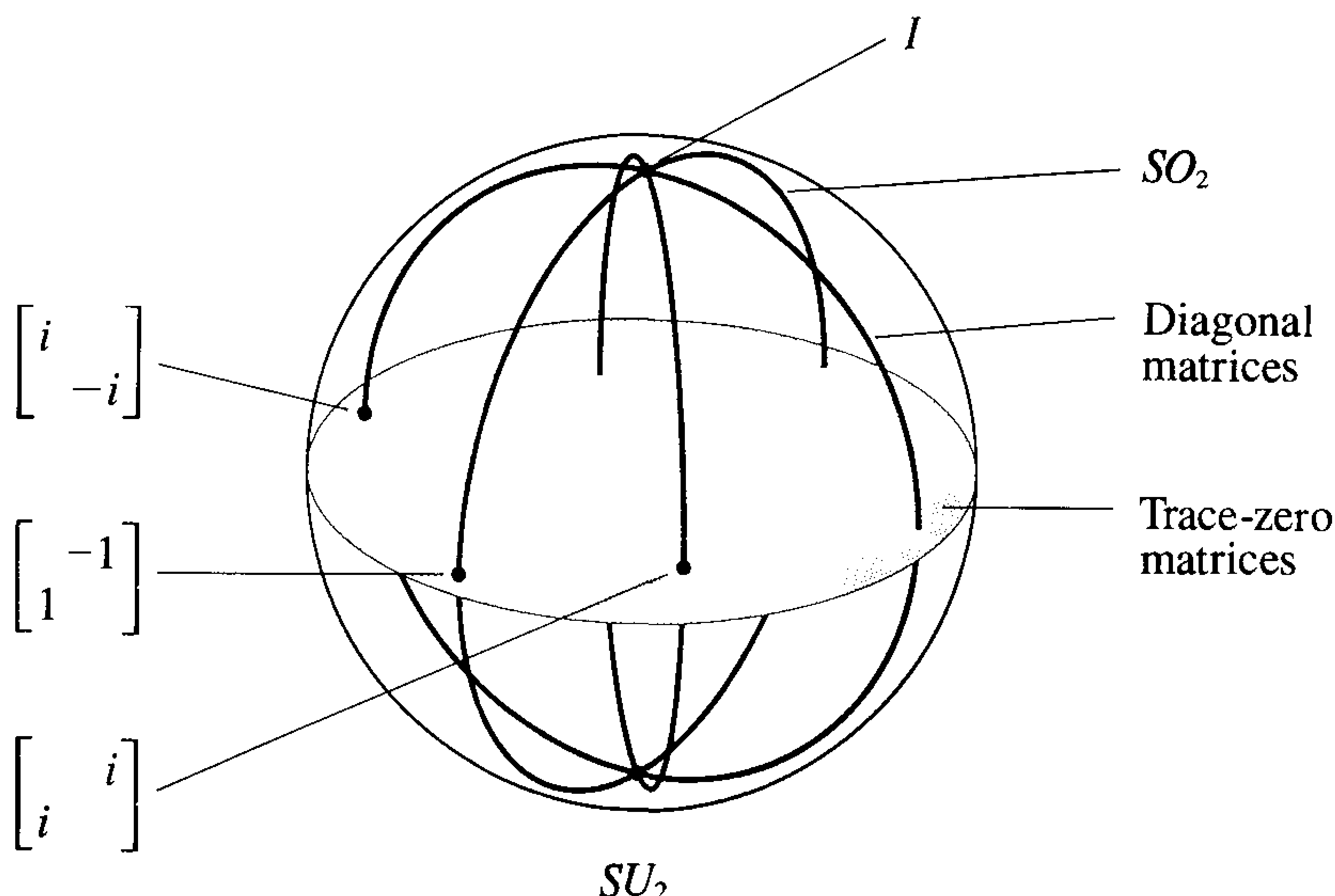
be linearly independent and thus will span a subspace V of $\mathbb{R}^4$ of dimension 2. The intersection $SU_2 \cap V$ is the unique longitude containing P.

The intersection of SU_2 with the plane W defined by $x_3 = x_4 = 0$ is a particularly nice longitude. In matrix notation, this great circle consists of the diagonal matrices in SU_2, which form a subgroup T:

(2.13)
$$T = \left\{ \begin{bmatrix} \lambda & \\ & \bar{\lambda} \end{bmatrix} \middle| \lambda\bar{\lambda} = 1 \right\}.$$

The other longitudes are described in the following proposition.

(2.14) **Proposition.** The longitudes of SU_2 are the conjugate subgroups QTQ^* of the subgroup T.



(2.15) **Figure.** Some latitudes and longitudes in SU_2.

In Figure (2.15) the 3-sphere SU_2 is projected from $\mathbb{R}^4$ onto the unit disc in the plane. The conjugacy class shown is the "equatorial" latitude in $\mathbb{R}^4$, which is defined by the equation $x_1 = 0$. Just as the orthogonal projection of a circle from $\mathbb{R}^3$ to $\mathbb{R}^2$ is an ellipse, the projection of this 2-sphere from $\mathbb{R}^4$ to $\mathbb{R}^3$ is an ellipsoid, and the further projection of this ellipsoid to the plane is the elliptical disc shown.

Proof of Proposition (2.14). The point here is to show that any conjugate subgroup QTQ^* is a longitude. Lemma (2.11) tells us that every element $P \in SU_2$ lies in one of these conjugate subgroups (though the roles of Q and Q^* have been reversed). Since every $P \neq \pm I$ is contained in exactly one longitude, it will follow that every longitude is one of the subgroups QTQ^*.

So let us show that a conjugate subgroup QTQ^* is a longitude. The reason this is true is that conjugation by a fixed element Q is a linear operator which sends the

subspace W to another subspace. We will compute the conjugate explicitly to make this clear. Say that Q is the matrix (2.4). Let $w = (w_1, w_2, 0, 0)$ denote a variable element of W, and set $z = w_1 + w_2 i$. Then

$$\begin{bmatrix} a & b \\ -\overline{b} & \overline{a} \end{bmatrix}\begin{bmatrix} z & \\ & \overline{z} \end{bmatrix}\begin{bmatrix} \overline{a} & -b \\ \overline{b} & a \end{bmatrix} = \begin{bmatrix} a\overline{a}z+b\overline{b}\overline{z} & ab(\overline{z}-z) \\ * & * \end{bmatrix}.$$

Computing these entries, we find that w is sent to the vector $u = (u_1, u_2, u_3, u_4)$, where

$$u_1 = w_1,\ u_2 = (x_1^2+x_2^2-x_3^2-x_4^2)w_2,$$
$$u_3 = 2(x_1x_4+x_2x_3)w_2,\ u_4 = 2(x_2x_4-x_1x_3)w_2.$$

The coordinates u_i are real linear combinations of (w_1, w_2). This shows that the map $w \rightsquigarrow u$ is a real linear transformation. So its image V is a subspace of $\mathbb{R}^4$. The conjugate group QTQ^* is $SU_2 \cap V$. Since QTQ^* contains the poles $\pm I$, so does V, and this shows that QTQ^* is a longitude. □

We will describe another geometric configuration briefly: As we have seen, the subgroup T of diagonal matrices is a great circle in the 3-sphere SU_2. The left cosets of this subgroup, the sets of the form QT for $Q \in SU_2$, are also great circles, and they partition the group SU_2. Thus the 3-sphere is partitioned into great circles. This very interesting configuration is called the *Hopf fibration.*

3. THE ORTHOGONAL REPRESENTATION OF SU_2

We saw in the last section that the conjugacy classes in the special unitary group SU_2 are two-dimensional spheres. Since conjugacy classes are orbits for the operation of conjugation, SU_2 operates on these spheres. In this section we will show that conjugation by an element $P \in SU_2$ acts on each of the spheres as a *rotation,* and that the map sending P to the matrix of this rotation defines a surjective homomorphism

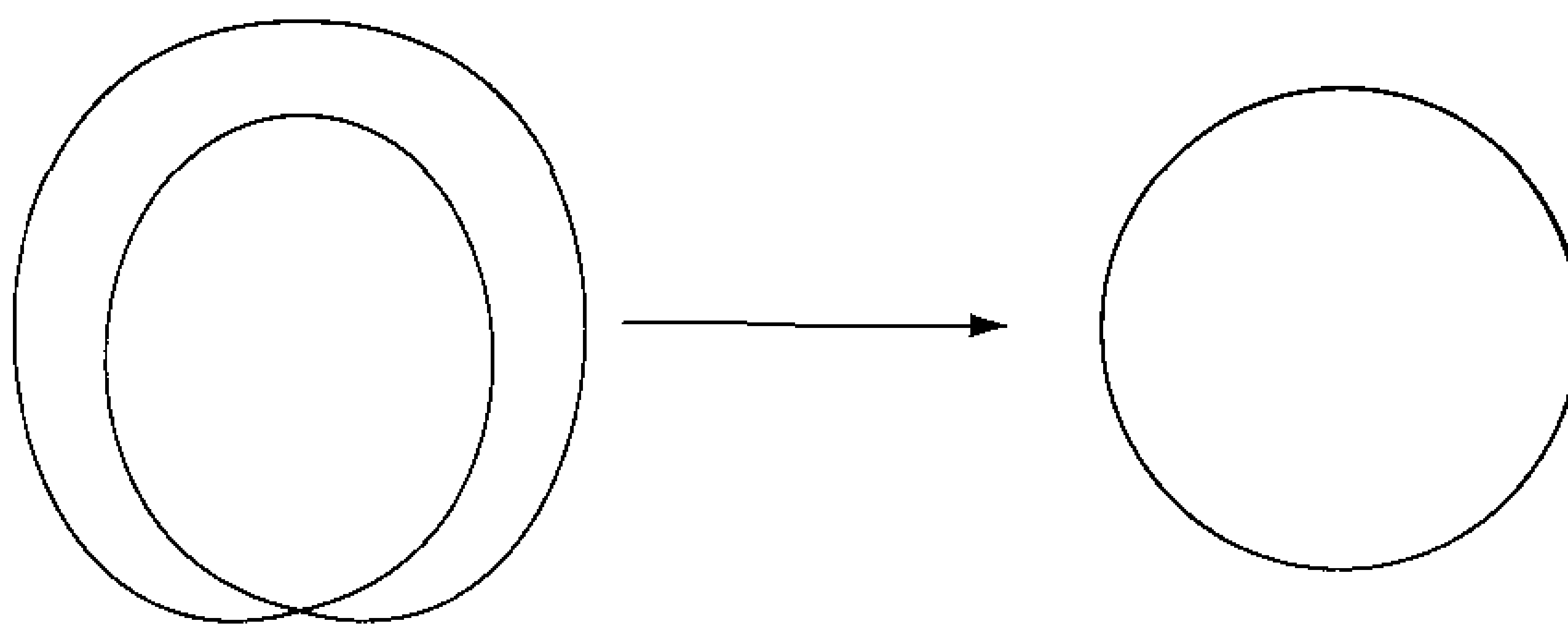
(3.1) $$\varphi: SU_2 \longrightarrow SO_3,$$

whose kernel is the center $Z = \{\pm I\}$ of SU_2. This homomorphism is called the *orthogonal representation* of SU_2. It represents a complex 2×2 matrix P in SU_2 by a real 3×3 rotation matrix $\varphi(P)$.

The safest way to show that P operates by rotating a conjugacy class may be to write the matrix representing the rotation down explicitly. This is done in (3.12). However, the formula for $\varphi(P)$ is complicated and not particularly enlightening. It is better to describe φ indirectly, as we will do presently. Let us discuss the geometry of the map first.

Since the kernel of φ is $\{\pm I\}$, its cosets are the sets $\{\pm P\}$. They form the fibres of the homomorphism. Thus every element of SO_3 corresponds to a pair of unitary matrices which differ by sign. Because of this, the group SU_2 is called a *double covering* of the group SO_3.

The map μ: $SO_2 \longrightarrow SO_2$ of the 1-sphere to itself defined by $\rho_\theta \rightsquigarrow \rho_{2\theta}$ is another example of a double covering. Its kernel also consists of two elements, the identity and the rotation by π. Every fibre of μ contains two rotations ρ_θ and $\rho_{\pi+\theta}$.



(3.2) **Figure.** A double covering of the 1-sphere.

The orthogonal representation can be used to identify the topological structure of the rotation group. In vector notation, if $P = (x_1, \ldots, x_4)$, then $-P = (-x_1, \ldots, -x_4)$, and the point $-P$ is called the *antipode* of P. So since points of the rotation group correspond to cosets $\{\pm P\}$, the group SO_3 can be obtained by identifying antipodal points on the 3-sphere SU_2. The space obtained in this way is called the *real projective 3-space:*

(3.3) *SO_3 is homeomorphic to the real projective 3-space.*

The number 3 refers again to the dimension of the space. Points of the real projective 3-space are also in bijective correspondence with lines through the origin (or one-dimensional subspaces) of $\mathbb{R}^4$. Every line through the origin meets the unit sphere in a pair of antipodal points.

As we noted in Section 8 of Chapter 4, every element of SO_3 except the identity can be described in terms of a pair (v, θ), where v is a unit vector in the axis of rotation and where θ is the angle of rotation. However, the two pairs (v, θ) and $(-v, -\theta)$ represent the same rotation. The choice of one of these pairs is referred to by physicists as the choice of a *spin*. It is not possible to make a choice of spin which varies continuously over the whole group. Instead, the two possible choices define a double covering of $SO_3 - \{I\}$. We may realize the set of all pairs (v, θ) as the product space $S \times \Theta$, where S is the 2-sphere of unit vectors in $\mathbb{R}^3$, and where Θ is the set of nonzero angles $0 < \theta < 2\pi$. This product space maps to SO_3:

(3.4) $$\psi: S \times \Theta \longrightarrow SO_3 - \{I\},$$

by sending (v, θ) to the rotation about v through the angle θ. The map ψ is a double covering of $SO_3 - \{I\}$ because every nontrivial rotation is associated to two pairs $(v, \theta), (-v, -\theta)$.

We now have two double coverings of $SO_3 - \{I\}$, namely $S \times \Theta$ and also $SU_2 - \{\pm I\}$, and it is plausible that they are equivalent. This is true:

(3.5) **Proposition.** There is a homeomorphism h: $(SU_2 - \{\pm I\}) \longrightarrow S \times \Theta$ which is compatible with the maps SO_3, i.e., such that $\psi \circ h = \varphi$.

This map h is not a group homomorphism. In fact, neither its domain nor its range is a group.

Proposition (3.5) is not very difficult to prove, but the proof is slightly elusive because there are two such homeomorphisms. They differ by a switch of the spin. On the other hand, the fact that this homeomorphism exists follows from a general theorem of topology, because the space $SU_2 - \{\pm I\}$ is *simply connected*. (A simply connected space is one which is path connected and such that every loop in the space can be contracted continuously to a point.) It is better to leave this proof to the topologists. □

Therefore every element of SU_2 except $\pm I$ can be described as a rotation of $\mathbb{R}^3$ together with a choice of spin. Because of this, SU_2 is often called the *Spin group*.

We now proceed to compute the homomorphism φ, and to begin, we must select a conjugacy class. It is convenient to choose the one consisting of the trace-zero matrices in SU_2, which is the one defined by $x_1 = 0$ and which is illustrated in Figure (2.15). The group operates in the same way on the other classes. Let us call the conjugacy class of trace-zero matrices C. An element A of C will be a matrix of the form

$$(3.6) \qquad A = \begin{bmatrix} y_2 i & y_3+y_4 i \\ -y_3+y_4 i & -y_2 i \end{bmatrix},$$

where

$$(3.7) \qquad y_2{}^2+y_3{}^2+y_4{}^2 = 1.$$

Notice that this matrix is *skew-hermitian,* that is, it has the property

$$(3.8) \qquad A^* = -A.$$

(We haven't run across skew-hermitian matrices before, but they aren't very different from hermitian matrices. In fact, A is a skew-hermitian matrix if and only if $H = iA$ is hermitian.) The 2×2 skew-hermitian matrices with trace zero form a real vector space V of dimension 3, with basis

$$(3.9) \qquad \mathbf{B} = \left(\begin{bmatrix} i & \\ & -i \end{bmatrix}, \begin{bmatrix} & 1 \\ -1 & \end{bmatrix}, \begin{bmatrix} & i \\ i & \end{bmatrix} \right).$$

In the notation of (3.6), $A = \mathbf{B}Y$, where $Y = (y_2, y_3, y_4)^t$. So the basis $\mathbf{B}$ corresponds to the standard basis (e_2, e_3, e_4) in the space $\mathbb{R}^3$, and (3.7) tells us that our conjugacy class is represented as the unit sphere in this space.

Note that SU_2 operates by conjugation on the whole space V of trace-zero, skew-hermitian matrices, not only on its unit sphere: If $A \in V$, $P \in SU_2$, and if $B = PAP^* = PAP^{-1}$, then trace $B = 0$, and $B^* = (PAP^*)^* = PA^*P^* = (P(-A)P^* = -B$. Also, conjugation by a fixed matrix P gives a linear operator on V, because $P(A + A')P^* = PAP^* + PA'P^*$, and if r is a real number, then $P(rA)P^* = rPAP^*$. The matrix of this linear operator is defined to be $\varphi(P)$. To determine the matrix ex-

plicitly, we conjugate the basis (3.7) by P and rewrite the result in terms of the basis. For example,

$$(3.10)\qquad \begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix}\begin{bmatrix} i & \\ & -i \end{bmatrix}\begin{bmatrix} \bar{a} & -b \\ \bar{b} & a \end{bmatrix} = i\begin{bmatrix} a\bar{a}-b\bar{b} & -2ab \\ -2\bar{a}\bar{b} & b\bar{b}-a\bar{a} \end{bmatrix}.$$

The coordinates of this matrix are $y_2 = a\bar{a} - b\bar{b}$, $y_3 = i(-ab + \bar{a}\bar{b})$, and $y_4 = -(ab + \bar{a}\bar{b})$. They form the first column of the matrix $\varphi(P)$. Similar computation for the other columns yields

$$(3.11)\qquad \begin{bmatrix} (a\bar{a}-b\bar{b}) & i(\bar{a}b-a\bar{b}) & (\bar{a}b+a\bar{b}) \\ i(\overline{ab}-ab) & \frac{1}{2}(a^2+\bar{a}^2+b^2+\bar{b}^2) & \frac{i}{2}(a^2-\bar{a}^2-b^2+\bar{b}^2) \\ -(\overline{ab}+ab) & \frac{i}{2}(\bar{a}^2-a^2+\bar{b}^2-b^2) & \frac{1}{2}(a^2+\bar{a}^2-b^2-\bar{b}^2) \end{bmatrix}.$$

We will not make use of the above computation. Even without it, we know that $\varphi(P)$ is a real 3×3 matrix because it is the matrix of a linear operator on a real vector space V of dimension 3.

(3.12) **Lemma.** The map $P \rightsquigarrow \varphi(P)$ defines a homomorphism $SU_2 \longrightarrow GL_3(\mathbb{R})$.

Proof. It follows from the associative law [Chapter 5 (5.1)] for the operation of conjugation that φ is compatible with multiplication: The operation of a product PQ on a matrix A is $(PQ)A(PQ)^* = P(QAQ^*)P^*$. This is the composition of the operations of conjugation by P and by Q. Since the matrix of the composition of linear operators is the product matrix, $\varphi(PQ) = \varphi(P)\varphi(Q)$. Being compatible with multiplication, $\varphi(P^{-1})\varphi(P) = \varphi(I_2) = I_3$. Therefore $\varphi(P)$ is invertible for every P, and so φ is a homomorphism from SU_2 to $GL_3(\mathbb{R})$, as asserted. □

(3.13) **Lemma.** For any P, $\varphi(P) \in SO_3$. Hence $P \rightsquigarrow \varphi(P)$ defines a homomorphism $SU_2 \longrightarrow SO_3$.

Proof. One could prove this lemma using Formula (3.11). To prove it conceptually, we note that dot product on $\mathbb{R}^3$ carries over to a bilinear form on V with a nice expression in terms of the matrices. Using the notation of (3.6), we define $\langle A, A' \rangle = y_1y_1' + y_2y_2' + y_3y_3'$. Then

$$(3.14)\qquad \langle A, A' \rangle = -\tfrac{1}{2}\operatorname{trace}(AA').$$

This is proved by computation:

$$AA' = \begin{bmatrix} -(y_2y_2'+y_3y_3'+y_4y_4') & * \\ * & -(y_2y_2'+y_3y_3'+y_3y_3') \end{bmatrix},$$

and so $\operatorname{tr} AA' = -2\langle A, A' \rangle$.

This expression for dot product shows that it is preserved by conjugation by an element $P \in SU_2$:

$$\langle PAP^*, PA'P^* \rangle = -\tfrac{1}{2}\operatorname{trace}(PAP^*PA'P^*) = -\tfrac{1}{2}\operatorname{trace}(AA') = \langle A, A' \rangle.$$

Or, in terms of the coordinate vectors, $(\varphi(P)Y \cdot \varphi(P)Y') = (Y \cdot Y')$. It follows that $\varphi(P)$ lies in the orthogonal group $O_3 = O_3(\mathbb{R})$ [Chapter 4 (5.13)].

To complete the proof, let us verify that $\varphi(P)$ has determinant 1 for every $P \in SU_2$: Being a sphere, SU_2 is path connected. So only one of the two possible values ± 1 can be taken on by the continuous function $\det \varphi(P)$. Since $\varphi(I_2) = I_3$ and $\det I_3 = 1$, the value is always $+1$, and $\varphi(P) \in SO_3$, as required. □

(3.15) Lemma. $\ker \varphi = \{\pm I\}$.

Proof. The kernel of φ consists of the matrices $P \in SU_2$ which act trivially on V, meaning that $PAP^* = A$ for all skew-hermitian matrices of trace zero. Suppose that P has the property $PAP^* = A$, or $PA = AP$, for all $P \in V$. We test it on the basis (3.7). The test leads to $b = 0$, $a = \overline{a}$, which gives the two possibilities $P = \pm I$, and they are in the kernel. So $\ker \varphi = \{\pm I\}$, as claimed. □

(3.16) Lemma. The image of the map φ is SO_3.

Proof. We first compute $\varphi(P)$ explicitly on the subgroup T of diagonal matrices in SU_2. Let $z = y_3 + y_4 i$. Then

$$(3.17) \qquad PAP^* = \begin{bmatrix} a & \\ & \overline{a} \end{bmatrix} \begin{bmatrix} y_2 i & z \\ -\overline{z} & -y_2 i \end{bmatrix} \begin{bmatrix} \overline{a} & \\ & a \end{bmatrix} = \begin{bmatrix} y_2 i & a^2 z \\ -\overline{a}^2 \overline{z} & -y_2 i \end{bmatrix}.$$

So $\varphi(P)$ fixes the first coordinate y_2 and it multiplies z by a^2. Since $|a| = 1$, we may write $a = e^{i\theta}$. Multiplication by $a^2 = e^{2i\theta}$ defines a rotation by 2θ of the complex z-plane. Therefore

$$(3.18) \qquad \varphi(P) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\theta & -\sin 2\theta \\ 0 & \sin 2\theta & \cos 2\theta \end{bmatrix}.$$

This shows that the image of φ in SO_3 contains the subgroup H of all rotations about the point $(1,0,0)^t$. This point corresponds to the matrix $E = \begin{bmatrix} i & \\ & -i \end{bmatrix}$. Since the unit sphere C is a conjugacy class, the operation of SU_2 is transitive. So if Y is any unit vector in $\mathbb{R}^3$, there is an element $Q \in SU_2$ such that $\varphi(Q)(1,0,0)^t = Y$, or in matrix notation, such that $QEQ^* = A$. The conjugate subgroup $\varphi(Q)H\varphi(Q)^*$ of rotations about Y is also in the image of φ. Since every element of SO_3 is a rotation, φ is surjective. □

The cosets making up the Hopf fibration which was mentioned at the end of last section, are the fibres of a continuous surjective map

$$(3.19) \qquad \pi: S^3 \longrightarrow S^2$$

from the 3-sphere to the 2-sphere. To define π, we interpret S^3 as the special unitary group SU_2, and S^2 as the conjugacy class C of trace-zero matrices, as above. We

set $E = \begin{bmatrix} i & \\ & -i \end{bmatrix}$, and we define $\pi(P) = PEP^*$, for $P \in SU_2$. The proof of the following proposition is left as an exercise.

(3.20) **Proposition.** The fibres of the map π are the left cosets QT of the group T of diagonal matrices in SU_2. □

4. THE SPECIAL LINEAR GROUP $SL_2(\mathbb{R})$

Since the special unitary group is a sphere, it is a compact set. As an example of a noncompact group, we will describe the special linear group $SL_2(\mathbb{R})$. To simplify notation, we denote $SL_2(\mathbb{R})$ by SL_2 in this section.

Invertible 2×2 matrices operate by left multiplication on the space $\mathbb{R}^2$ of column vectors, and we can look at the associated action on *rays* in $\mathbb{R}^2$. A ray is a half line $R = \{rX \mid r \geq 0\}$. The set of rays is in bijective correspondence with the points on the unit circle S^1, the ray R corresponding to the point $R \cap S^1$.

Our group SL_2 operates by left multiplication on the set of rays. Let us denote by H the stabilizer of the ray $R_1 = \{re_1\}$ in $SL_2(\mathbb{R})$. It consists of matrices

(4.1)
$$B = \begin{bmatrix} a & b \\ 0 & a^{-1} \end{bmatrix},$$

where a is positive and b is arbitrary.

The rotation group SO_2 is another subgroup of SL_2, and it operates transitively on the set of rays.

(4.2) **Proposition.** The map $f: SO_2 \times H \longrightarrow SL_2$ defined by $f(Q,B) = QB$ is a homeomorphism (but not a group homomorphism).

Proof. Notice that $H \cap SO_2 = \{I\}$. Therefore f is injective [Chapter 2 (8.6)]. To prove surjectivity of f, let P be an arbitrary element of SL_2, and let R_1 be the ray $\{re_1 \mid r > 0\}$. Choose a rotation $Q \in SO_2$ such that $PR_1 = QR_1$. Then $Q^{-1}P$ is in the stabilizer H, say $Q^{-1}P = B$, or

(4.3)
$$P = QB.$$

Since f is defined by matrix multiplication, it is a continuous map. Also, in the construction of the inverse map, the rotation Q depends continuously on P because the ray PR_1 does. Then $B = Q^{-1}P$ also is a continuous function of P, and this shows that f^{-1} is continuous as well. □

Note that H can be identified by the rule $B \longleftrightarrow (a, b)$ with the product space (*positive reals*) $\times\ \mathbb{R}$. And the space of positive reals is homeomorphic by the log

function to the space $\mathbb{R}$ of all real numbers. Thus H is homeomorphic to $\mathbb{R}^2$. Since SO_2 is a circle, we find that

(4.4) *$SL_2(\mathbb{R})$ is homeomorphic to the product space $S^1 \times \mathbb{R}^2$.*

The special linear group can be related to the Lorentz group $O_{2,1}$ of two-dimensional space-time by a method analogous to that used in Section 3 for the orthogonal representation of SU_2. Let the coordinates in $\mathbb{R}^3$ be y_1, y_2, t, with the Lorentz form

(4.5)
$$y_1y_1' + y_2y_2' - tt',$$

and let W be the space of real trace-zero matrices. Using the basis

(4.6)
$$\begin{bmatrix} 1 & \\ & -1 \end{bmatrix}, \begin{bmatrix} & 1 \\ 1 & \end{bmatrix}, \begin{bmatrix} & 1 \\ -1 & \end{bmatrix},$$

we associate to a coordinate vector $(y_1, y_2, t)^t$ the matrix

(4.7)
$$A = \begin{bmatrix} y_1 & y_2+t \\ y_2-t & -y_1 \end{bmatrix}.$$

We use this representation of trace-zero matrices because the Lorentz form (4.5) has a simple matrix interpretation on such matrices:

(4.8)
$$\langle A, A' \rangle = y_1y_1' + y_2y_2' - tt' = \tfrac{1}{2}\operatorname{trace}(AA').$$

The group SL_2 acts on W by conjugation,

(4.9)
$$P, A \rightsquigarrow PAP^{-1},$$

and this action preserves the Lorentz form on W, because

$$\operatorname{trace}(AA') = \operatorname{trace}((PAP^{-1})(PA'P^{-1})),$$

as in the previous section. Since conjugation is a linear operator on W, it defines a homomorphism $\varphi: SL_2 \longrightarrow GL_3(\mathbb{R})$. Since conjugation preserves the Lorentz form, the image $\varphi(P)$ of P is an element of $O_{2,1}$.

(4.10) **Theorem.** The kernel of the homomorphism φ is the subgroup $\{\pm I\}$, and the image is the path-connected component $O_{2,1}{}^0$ of $O_{2,1}$ containing the identity I. Therefore $O_{2,1}{}^0 \approx SL_2(\mathbb{R})/\{\pm I\}$.

It can be shown that the two-dimensional Lorentz group has four path-connected components.

The fact that the kernel of φ is $\{\pm I\}$ is easy to check, and the last assertion of the theorem follows from the others. We omit the proof that the image of φ is the subgroup $O_{2,1}{}^0$. □

5. ONE-PARAMETER SUBGROUPS

In Chapter 4, we defined the exponential of a matrix by the series

$$e^A = I + (1/1!)A + (1/2!)A^2 + (1/3!)A^3 + \cdots. \tag{5.1}$$

We will now use this function to describe the homomorphisms from the additive group of real numbers to the general linear group, which are differentiable functions of the variable $t \in \mathbb{R}$. Such a homomorphism is called a *one-parameter subgroup* of GL_n. (Actually, this use of the phrase "one-parameter subgroup" to describe such homomorphisms is a misnomer. The image of φ should be called the subgroup.)

(5.2) **Proposition.**

(a) Let A be an arbitrary real or complex matrix, and let GL_n denote $GL_n(\mathbb{R})$ *or* $GL_n(\mathbb{C})$, according to the case. The map $\varphi: \mathbb{R}^+ \longrightarrow GL_n$ defined by $\varphi(t) = e^{tA}$ is a group homomorphism.

(b) Conversely, let $\varphi: \mathbb{R}^+ \longrightarrow GL_n$ be a homomorphism which is a differentiable function of the variable $t \in \mathbb{R}^+$, and let A denote its derivative $\varphi'(0)$ at the origin. Then $\varphi(t) = e^{tA}$ for all t.

Proof. For any two real numbers r, s, the two matrices rA and sA commute. So Chapter 4 (7.13) tells us that

$$e^{(r+s)A} = e^{rA}e^{sA}. \tag{5.3}$$

This shows that $\varphi(t) = e^{tA}$ is a homomorphism. Conversely, let φ be a differentiable homomorphism $\mathbb{R}^+ \longrightarrow GL_n$. The assumption that φ is a homomorphism allows us to compute its derivative at any point. Namely, it tells us that $\varphi(t + \Delta t) = \varphi(\Delta t)\varphi(t)$ and $\varphi(t) = \varphi(0)\varphi(t)$. Thus

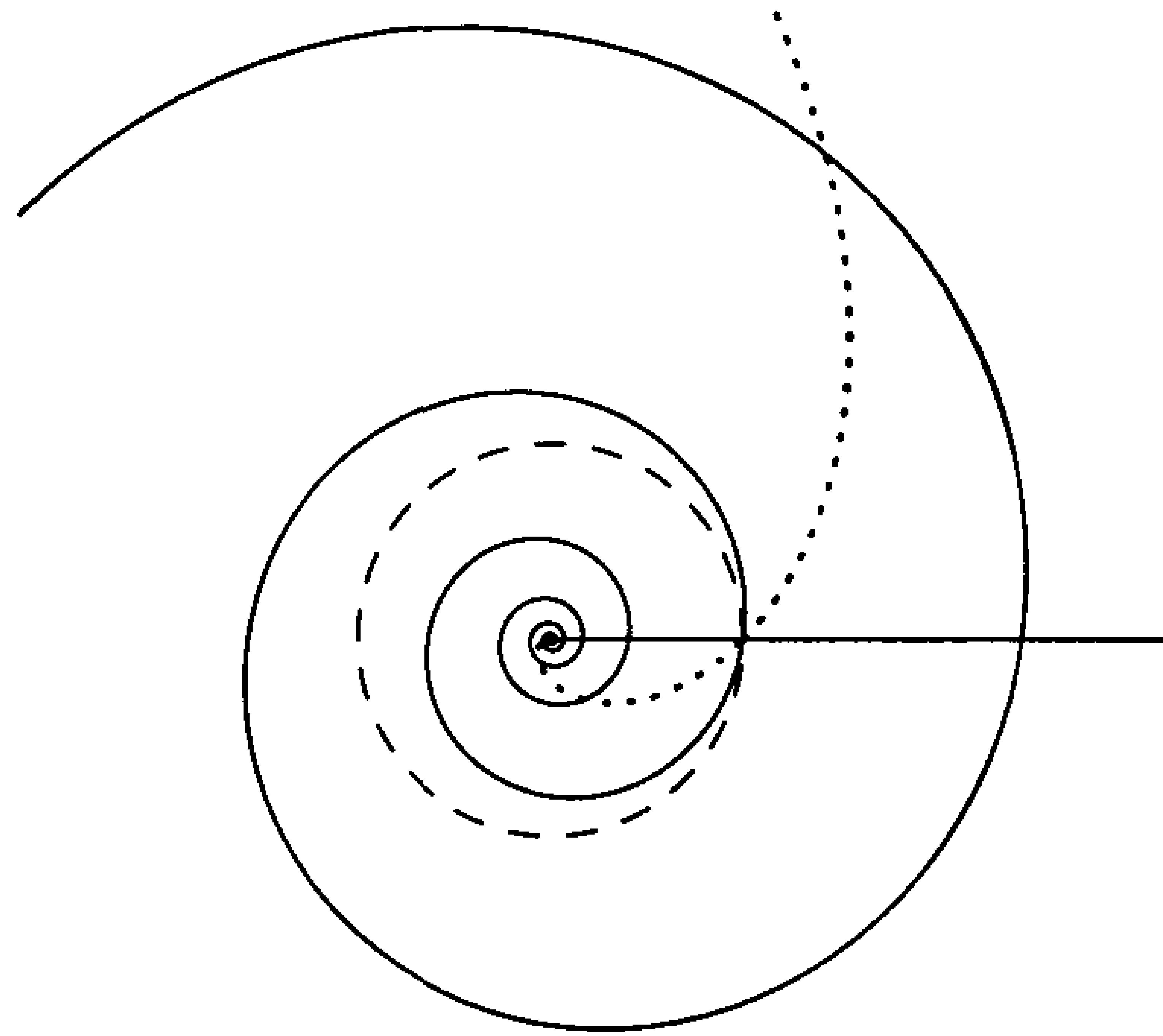
$$\frac{\varphi(t + \Delta t) - \varphi(t)}{\Delta t} = \frac{\varphi(\Delta t) - \varphi(0)}{\Delta t}\varphi(t). \tag{5.4}$$

Letting $\Delta t \longrightarrow 0$, we find $\varphi'(t) = \varphi'(0)\varphi(t) = A\varphi(t)$. Therefore $\varphi(t)$ is a matrix-valued function which solves the differential equation

$$\frac{d\varphi}{dt} = A\varphi. \tag{5.5}$$

The function e^{tA} is another solution, and both solutions take the value I at $t = 0$. It follows that $\varphi(t) = e^{tA}$ [see Chapter 4 (8.14)]. □

By the proposition we have just proved, the one-parameter subgroups all have the form $\varphi(t) = e^{tA}$. They are in bijective correspondence with $n \times n$ matrices.



(5.6) **Figure.** Some one-parameter subgroups of $\mathbb{C}^\times = GL_1(\mathbb{C})$.

Now suppose that a subgroup of G of GL_n is given. We may ask for one-parameter subgroups of G, meaning homomorphisms $\varphi: \mathbb{R}^+ \to G$, or, equivalently, homomorphisms to GL_n whose image is in G. Since a one-parameter subgroup of GL_n is determined by a matrix, this amounts to asking for the matrices A such that $e^{tA} \in G$ for all t. It turns out that linear groups of positive dimension always have one-parameter subgroups and that they are not hard to determine for a particular group.

(5.7) **Examples.**

(a) The usual parametrization of the unit circle in the complex plane is a one-parameter subgroup of U_1:

$$t \rightsquigarrow e^{ti} = \cos t + i \sin t .$$

(b) A related example is obtained for SO_2 by setting

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \text{ Then } e^{tA} = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}.$$

This is the standard parametrization of the rotation matrices.

In examples (a) and (b), the image of the homomorphism is the whole subgroup.

(c) Let A be the 2×2 matrix unit e_{12}. Then since $A^2 = 0$, all but two terms of the series expansion for the exponential vanish, and

$$e^{tA} = I + e_{12}t = \begin{bmatrix} 1 & t \\ & 1 \end{bmatrix}.$$

In this case the exponential map defines an isomorphism from $\mathbb{R}^+$ to its image, which is the group of triangular matrices with diagonal entries equal to 1.

(d) The one-parameter subgroups of SU_2 are the conjugates of the group of diagonal special unitary matrices, the longitudes described in (2.13). □

Instead of attempting to state a general theorem describing one-parameter subgroups of a group, we will determine them for the orthogonal and special linear groups as examples of the methods used. We will need to know that the exponential function on matrices has an inverse function.

(5.8) **Proposition.** The matrix exponential maps a small neighborhood S of 0 in $\mathbb{R}^{n\times n}$ homeomorphically to a neighborhood T of I.

Proof. This proposition follows from the Inverse Function Theorem, which states that a differentiable function $f: \mathbb{R}^k \longrightarrow \mathbb{R}^k$ has an inverse function at a point p if the Jacobian matrix $(\partial f_i/\partial x_j)(p)$ is invertible. We must check this for the matrix exponential at the zero matrix in $\mathbb{R}^{n\times n}$. This is a notationally unpleasant but easy computation. Let us denote a variable matrix by X. The Jacobian matrix is the $n^2 \times n^2$ matrix whose entries are $(\partial(e^X)_{\alpha\beta}/\partial X_{ij})|_{x=0}$. We use the fact that $d/dt\,(e^{tA})|_{t=0} = A$. It follows directly from the definition of the partial derivative that $(\partial e^X/\partial X_{ij})|_{X=0} = (de^{te_{ij}}/dt)|_{t=0} = e_{ij}$. Therefore $(\partial(e^X)_{\alpha\beta}/\partial X_{ij})|_{x=0} = 0$ if $\alpha, \beta \neq i, j$ and $(\partial(e^X)_{ij}/\partial X_{ij})|_{x=0} = 1$. The Jacobian matrix is the $n^2 \times n^2$ identity matrix. □

We will now describe one-parameter subgroups of the orthogonal group O_n. Here we are asking for the matrices A such that e^{tA} is orthogonal for all t.

(5.9) **Lemma.** If A is skew-symmetric, then e^A is orthogonal. Conversely, there is a neighborhood S' of 0 in $\mathbb{R}^{n\times n}$ such that if e^A is orthogonal and $A \in S'$, then A is skew-symmetric.

Proof. To avoid confusing the variable t with the symbol for the transpose matrix, we denote the transpose of the matrix A by A^* here. If A is skew-symmetric, then $e^{(A^*)} = e^{-A}$. The relation $e^{(A^*)} = (e^A)^*$ is clear from the definition of the exponential, and $e^{-A} = (e^A)^{-1}$ by Chapter 4 (8.10). Thus $(e^A)^* = e^{(A^*)} = e^{-A} = (e^A)^{-1}$. This shows that e^A is orthogonal. For the converse, we choose S' small enough so that if $A \in S'$, then $-A$ and A^* are in the neighborhood S of Proposition (5.8). Suppose that $A \in S'$ and that e^A is orthogonal. Then $e^{(A^*)} = e^{-A}$, and by Proposition (5.8), this means that A is skew-symmetric. □

(5.10) **Corollary.** The one-parameter subgroups of the orthogonal group O_n are the homomorphisms $t \rightsquigarrow e^{tA}$, where A is a real skew-symmetric matrix.

Proof. If A is skew-symmetric, tA is skew-symmetric for all t. So e^{tA} is orthogonal for all t, which means that e^{tA} is a one-parameter subgroup of O_n. Conversely, suppose that e^{tA} is orthogonal for all t. For sufficiently small ϵ, ϵA is in the neighborhood S' of the lemma, and $e^{\epsilon A}$ is orthogonal. Therefore ϵA is skew-symmetric, and this implies that A is skew-symmetric too. □

This corollary is illustrated by Example (5.7b).

Next, let us consider the special linear group $SL_n(\mathbb{R})$.

(5.11) **Proposition.** Let A be a matrix whose trace is zero. Then e^A has determinant 1. Conversely, there is a neighborhood S' of 0 in $\mathbb{R}^{n\times n}$ such that if $A \in S'$ and $\det e^A = 1$, then trace $A = 0$.

Proof. The first assertion follows from the pretty formula

$$e^{\operatorname{tr} A} = \det e^A, \tag{5.12}$$

where tr$A$ denotes the trace of the matrix. This formula follows in turn from the fact that if the eigenvalues of a complex matrix A are $\lambda_1,\dots,\lambda_n$, then the eigenvalues of e^A are $e^{\lambda_1},\dots,e^{\lambda_n}$. We leave the proof of this fact as an exercise. Using it, we find

$$e^{\operatorname{tr} A} = e^{\lambda_1+\cdots+\lambda_n} = e^{\lambda_1}\cdots e^{\lambda_n} = \det e^A.$$

For the converse, we note that if $|x| < 1$, $e^x = 1$ implies $x = 0$. We choose S' small enough so that tr $A < 1$ if $A \in S'$. Then if $\det e^A = e^{\operatorname{tr} A} = 1$ and if $A \in S'$, tr $A = 0$. □

(5.13) **Corollary.** The one-parameter subgroups of the special linear group $SL_n(\mathbb{R})$ are the homomorphisms $t \rightsquigarrow e^{tA}$, where A is a real $n \times n$ matrix whose trace is zero. □

The simplest one-parameter subgroup of $SL_2(\mathbb{R})$ is described in Example (5.7c).

6. THE LIE ALGEBRA

As always, we think of a linear group G as a subset of $\mathbb{R}^{n\times n}$ or of $\mathbb{C}^{n\times n}$. The space of vectors tangent to G at the identity matrix I, which we will describe in this section, is called the *Lie algebra* of the group.

We will begin by reviewing the definition of tangent vector. If $\varphi(t) = (\varphi_1(t),\dots,\varphi_k(t))$ is a differentiable path in $\mathbb{R}^k$, its velocity vector $v = \varphi'(t)$ is tangent to the path at the point $x = \varphi(t)$. This is the basic observation from which the definition of tangent vector is derived.

Suppose that we are given a subset S of $\mathbb{R}^k$. A vector v is said to be *tangent* to S at a point x if there is a differentiable path $\varphi(t)$ lying entirely in S, such that $\varphi(0) = x$ and $\varphi'(0) = v$.

If our subset S is the locus of zeros of one or more polynomial functions $f(x_1,\dots,x_k)$, it is called a *real algebraic set:*

$$S = \{x \mid f(x) = 0\}. \tag{6.1}$$

For example, the unit circle in $\mathbb{R}^2$ is a real algebraic set because it is the locus of zeros of the polynomial $f(x_1, x_2) = x_1^2 + x_2^2 - 1 = 0$.

The chain rule for differentiation provides a necessary condition for a vector to be tangent to a real algebraic set S. Let $\varphi(t)$ be a path in S, and let $x = \varphi(t)$ and

$v = \varphi'(t)$. Since the path is in S, the functions $f(\varphi(t))$ vanish identically; hence their derivatives also vanish identically:

$$(6.2) \qquad 0 = \frac{d}{dt} f(\varphi(t)) = \frac{\partial f}{\partial x_1} v_1 + \cdots + \frac{\partial f}{\partial x_k} v_k = (\nabla f(x) \cdot v),$$

where $\nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_k} \right)$ is the gradient vector.

(6.3) **Corollary.** Let S be a real algebraic set in $\mathbb{R}^k$, defined as the locus of zeros of one or more polynomial functions $f(x)$. The tangent vectors to S at x are orthogonal to the gradients $\nabla f(x)$. □

For instance, if S is the unit circle and x is the point $(1, 0)$, then the gradient vector $\nabla f(0)$ is $(2, 0)$. Corollary (6.3) tells us that tangent vectors at $(1, 0)$ have the form $(0, c)$, that is, that they are vertical, which is as it should be.

Computing tangent vectors by means of parametrized paths is clumsy because there are many paths with the same tangent. If we are interested only in the tangent vector, then we can throw out all of the information contained in a path except for the first-order term of its Taylor expansion. To do this systematically, we introduce a formal *infinitesimal element* ϵ. This means that we work algebraically with the rule

$$(6.4) \qquad \epsilon^2 = 0.$$

Just as with complex numbers, where the rule is $i^2 = -1$, we can use this rule to define a multiplication on the vector space

$$E = \{a + b\epsilon \mid a, b \in \mathbb{R}\}$$

of formal linear combinations of $(1, \epsilon)$ with real coefficients. The rule for multiplication is

$$(6.5) \qquad (a + b\epsilon)(c + d\epsilon) = ac + (bc + ad)\epsilon.$$

In other words, we expand formally, using the relations $\epsilon c = c\epsilon$ for all $c \in \mathbb{R}$ and $\epsilon^2 = 0$. As with complex numbers, addition is vector addition:

$$(a + b\epsilon) + (c + d\epsilon) = (a + c) + (b + d)\epsilon.$$

The main difference between $\mathbb{C}$ and E is that E is not a field, because ϵ has no multiplicative inverse. [It is a ring (see Chapter 10).]

Given a point x of $\mathbb{R}^k$ and a vector $v \in \mathbb{R}^k$, the sum $x + v\epsilon$ is a vector with entries in E which we interpret intuitively as an *infinitesimal change in x, in the direction of v*. Notice that we can evaluate a polynomial $f(x) = f(x_1, \ldots, x_k)$ at $x + v\epsilon$ using Taylor's expansion. Since $\epsilon^2 = 0$, the terms of degree ≥ 2 in ϵ drop out, and we are left with an element of E:

$$(6.6) \quad f(x + v\epsilon) = f(x) + \left(\frac{\partial f}{\partial x_1} v_1 + \cdots + \frac{\partial f}{\partial x_k} v_k \right) \epsilon = f(x) + (\nabla f(x) \cdot v)\epsilon.$$

Working with rule (6.4) amounts to ignoring the higher-order terms in ϵ. Thus the dot product $(\nabla f(x) \cdot v)$ represents the infinitesimal change in f which results when we make an infinitesimal change in x in the direction of v.

Going back to a real algebraic set S defined by the polynomial equations $f(x) = 0$, let x be a point of S. Then $f(x) = 0$, so (6.6) tells us that

$$f(x + v\epsilon) = 0 \text{ if and only if } (\nabla f(x) \cdot v) = 0, \tag{6.7}$$

which is the same as the condition we obtained in Corollary (6.3). This suggests the following definition: Let S be a real algebraic set, defined by the polynomial equations $f(x) = 0$. A vector v is called an *infinitesimal tangent* to S at x if

$$f(x + v\epsilon) = 0. \tag{6.8}$$

(6.9) **Corollary.** Let x be a point of a real algebraic set S. Every tangent to S at x is an infinitesimal tangent. □

Notice that if we fix $x \in S$, the equations $(\nabla f(x) \cdot v) = 0$ are linear and homogeneous in v. So the infinitesimal tangent vectors to S at x form a subspace of the space of all vectors.

Actually, our terminology is slightly ambiguous. The definition of an infinitesimal tangent depends on the equations f, not only on the set S. *We must have particular equations in mind when speaking of infinitesimal tangents.*

For sets S which are sufficiently smooth, the converse of (6.9) is also true: Every infinitesimal tangent is a tangent vector. When this is the case, we can compute the space of tangent vectors at a point $x \in S$ by solving the linear equations $(\nabla f(x) \cdot v) = 0$ for v, which is relatively easy. However, this converse will not be true at "singular points" of the set S, or if the defining equations for S are chosen poorly. For example, let S denote the union of the two coordinate axes in $\mathbb{R}^2$. This is a real algebraic set defined by the single equation $x_1x_2 = 0$. It is clear that at the origin a tangent vector must be parallel to one of the two axes. On the other hand, $\nabla f = (x_2, x_1)$, which is zero when $x_1 = x_2 = 0$. Therefore *every* vector is an infinitesimal tangent to S at the origin.

This completes our general discussion of tangent vectors. We will now apply this discussion to the case that the set S is one of our linear groups G in $\mathbb{R}^{n \times n}$ or $\mathbb{C}^{n \times n}$. The tangent vectors to G will be n^2-dimensional vectors, and we will represent them by matrices too. As we said earlier, the vectors tangent to G at the identity I form the *Lie algebra* of the group.

The first thing to notice is that every one-parameter subgroup e^{tA} of our linear group G is a parametrized path. We already know that its velocity vector $(de^{tA}/dt)_{t=0}$ is A. So A represents a tangent vector to G at the identity—it is in the Lie algebra. For example, the unitary group U_1 is the unit circle in the complex plane, and e^{ti} is a one-parameter subgroup of U_1. The velocity vector of this one-parameter subgroup at $t = 0$ is the vector i, which is indeed a tangent vector to the unit circle at the point 1.

A matrix group G which is a real algebraic set in $\mathbb{R}^{n\times n}$ is called a *real algebraic group*. The classical linear groups such as $SL_n(\mathbb{R})$ and O_n are real algebraic, because their defining equations are polynomial equations in the matrix entries. For example, the group $SL_2(\mathbb{R})$ is defined by the single polynomial equation $\det P = 1$:

$$x_{11}x_{22} - x_{12}x_{21} - 1 = 0.$$

The orthogonal group O_3 is defined by nine polynomials f_{ij} expressing the condition $P^t P = I$:

$$f_{ij} = x_{1i}x_{1j} + x_{2i}x_{2j} + x_{3i}x_{3j} - \delta_{ij} = 0, \quad \delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}.$$

Complex groups such as the unitary groups can also be made into real algebraic groups in $\mathbb{R}^{2n\times n}$ by separating the matrix entries into their real and imaginary parts.

It is a fact that for every infinitesimal tangent A to a real algebraic group G at the identity, e^{tA} is a one-parameter subgroup of G. In other words, there is a one-parameter subgroup leading out from the identity in an arbitrary tangent direction. This is quite remarkable for a nonabelian group, but it is true with essentially no restriction. Unfortunately, though this fact is rather easy to check for a particular group, it is fairly hard to give a general proof. Therefore we will content ourselves with verifying particular cases.

Having an infinitesimal element available, we may work with matrices whose entries are in E. Such a matrix will have the form $A + B\epsilon$, where A,B are real matrices. Intuitively, $A + B\epsilon$ represents an infinitesimal change in A in the direction of the matrix B. The rule for multiplying two such matrices is the same as (6.5):

$$(6.10) \qquad (A + B\epsilon)(C + D\epsilon) = AC + (AD + BC)\epsilon.$$

The product $B\epsilon D\epsilon$ is zero because $(b_{ij}\epsilon)(d_{kl}\epsilon) = 0$ for all values of the indices.

Let G be a real algebraic group. To determine its infinitesimal tangent vectors at the identity, we must determine the matrices A such that

$$(6.11) \qquad I + A\epsilon,$$

which represents an infinitesimal change in I in the direction of the matrix A, satisfies the equations defining G. This is the definition (6.8) of an infinitesimal tangent.

Let us make this computation for the special linear group $SL_n(\mathbb{R})$. The defining equation for this group is $\det P = 1$. So A is an infinitesimal tangent vector if $\det(I + A\epsilon) = 1$. To describe this condition, we must calculate the change in the determinant when we make an infinitesimal change in I. The formula is nice:

$$(6.12) \qquad \det(I + A\epsilon) = 1 + (\operatorname{trace} A)\epsilon.$$

The proof of this formula is left as an exercise. Using it, we find that A is an infinitesimal tangent vector if and only if $\operatorname{trace} A = 0$.

(6.13) **Proposition.** The following conditions on a real $n \times n$ matrix A are equivalent:

(i) $\operatorname{trace} A = 0$;
(ii) e^{tA} is a one-parameter subgroup of $SL_n(\mathbb{R})$;
(iii) A is in the Lie algebra of $SL_n(\mathbb{R})$;
(iv) A is an infinitesimal tangent to $SL_n(\mathbb{R})$ at I.

Proof. Proposition (5.11) tells us that (i) $\Rightarrow$ (ii). Since A is tangent to the path e^{tA} at $t = 0$, (ii) $\Rightarrow$ (iii). The implication (iii) $\Rightarrow$ (iv) is (6.9), and (iv) $\Rightarrow$ (i) follows from (6.12). □

There is a general principle at work here. We have three sets of matrices A: those such that e^{tA} is a one-parameter subgroup of G, those which are in the Lie algebra, and those which are infinitesimal tangents. Let us denote these three sets by $\operatorname{Exp}(G)$, $\operatorname{Lie}(G)$, and $\operatorname{Inf}(G)$. They are related by the following inclusions:

(6.14) $$\operatorname{Exp}(G) \subset \operatorname{Lie}(G) \subset \operatorname{Inf}(G).$$

The first inclusion is true because A is the tangent vector to e^{tA} at $t = 0$, and the second holds because every tangent vector is an infinitesimal tangent. If $\operatorname{Exp}(G) = \operatorname{Inf}(G)$, then these two sets are also equal to $\operatorname{Lie}(G)$. Since the computations of $\operatorname{Exp}(G)$ and $\operatorname{Inf}(G)$ are easy, this gives us a practical way of determining the Lie algebra. A general theorem exists which implies that $\operatorname{Exp}(G) = \operatorname{Inf}(G)$ for every real algebraic group, provided that its defining equations are chosen properly. However, it isn't worthwhile proving the general theorem here.

We will now make the computation for the orthogonal group O_n. The defining equation for O_n is the matrix equation $P^tP = I$. In order for A to be an infinitesimal tangent at the identity, it must satisfy the relation

(6.15) $$(I + A\epsilon)^t(I + A\epsilon) = I.$$

The left side of this relation expands to $I + (A^t + A)\epsilon$, so the condition that $I + A\epsilon$ be orthogonal is $A^t + A = 0$, or A is skew-symmetric. This agrees with the condition (5.10) for e^{tA} to be a one-parameter subgroup of O_n.

(6.16) **Proposition.** The following conditions on a real $n \times n$ matrix A are equivalent:

(i) A is skew-symmetric;
(ii) e^{tA} is a one-parameter subgroup of O_n;
(iii) A is in the Lie algebra of O_n;
(iv) A is an infinitesimal tangent to O_n at I. □

The Lie algebra of a linear group has an additional structure, an operation called the *Lie bracket*. The Lie bracket is the law of composition defined by the rule

(6.17) $$[A,B] = AB - BA.$$

This law of composition is not associative. It does, however, satisfy an identity called the *Jacobi identity*,

$$(6.18) \qquad [A,[B,C]] + [B,[C,A]] + [C,[A,B]] = 0,$$

which is a substitute for the associative law.

To show that the bracket is a law of composition on the Lie algebra, we must check that if A, B are in Lie(G), then $[A, B]$ is also in Lie(G). This can be done easily for any particular group. For the special linear group, the required verification is that if A, B have trace zero, then $AB - BA$ also has trace zero. This is true, because trace AB = trace BA. Or let $G = O_n$, so that the Lie algebra is the space of skew-symmetric matrices. We must verify that if A, B are skew, then $[A, B]$ is skew too:

$$[A,B]^t = (AB - BA)^t = B^tA^t - A^tB^t = BA - AB = -[A,B],$$

as required.

The bracket operation is important because it is the infinitesimal version of the commutator $PQP^{-1}Q^{-1}$. To see why this is so, we must work with two infinitesimals ϵ, δ, using the rules $\epsilon^2 = \delta^2 = 0$ and $\epsilon\delta = \delta\epsilon$. Note that the inverse of the matrix $I + A\epsilon$ is $I - A\epsilon$. So if $P = I + A\epsilon$ and $Q = I + B\delta$, the commutator expands to

$$(6.19) \qquad (I + A\epsilon)(I + B\delta)(I - A\epsilon)(I - B\delta) = I + (AB - BA)\epsilon\delta.$$

Intuitively, the bracket is in the Lie algebra because the product of two elements in G, even infinitesimal ones, is in G, and therefore the commutator of two elements is also in G.

Using the bracket operation, we can also define the concept of Lie algebra abstractly.

(6.20) **Definition.** A *Lie algebra* V over a field F is a vector space together with a law of composition

$$V \times V \longrightarrow V$$
$$v, w \rightsquigarrow [v, w]$$

called the *bracket*, having these properties:

(i) bilinearity: $[v_1 + v_2, w] = [v_1, w] + [v_2, w]$, $[cv, w] = c[v, w]$,
$[v, w_1 + w_2] = [v, w_1] + [v, w_2]$, $[v, cw] = c[v, w]$,

(ii) skew symmetry: $[v, v] = 0$,

(iii) Jacobi identity: $[u,[v,w]] + [v,[w,u]] + [w,[u,v]] = 0$,

for all $u, v, w \in V$ and all $c \in F$.

The importance of Lie algebras comes from the fact that, being vector spaces, they are much easier to work with than the linear groups themselves, and at the same time the classical groups are nearly determined by their Lie algebras. In other

words, the infinitesimal structure of the group at the identity element is almost enough to determine the group.

7. TRANSLATION IN A GROUP

We will use one more notion from topology in this section—the definition of manifold in $\mathbb{R}^k$. This definition is reviewed in the appendix [Definition (3.12)]. Do not be discouraged if you are not familiar with the concept of manifold. You can learn what is necessary without much trouble as we go along.

Let P be a fixed element of a matrix group G. We know that left multiplication by P is a bijective map from G to itself:

$$G \xrightarrow{m_P} G \qquad X \rightsquigarrow PX, \tag{7.1}$$

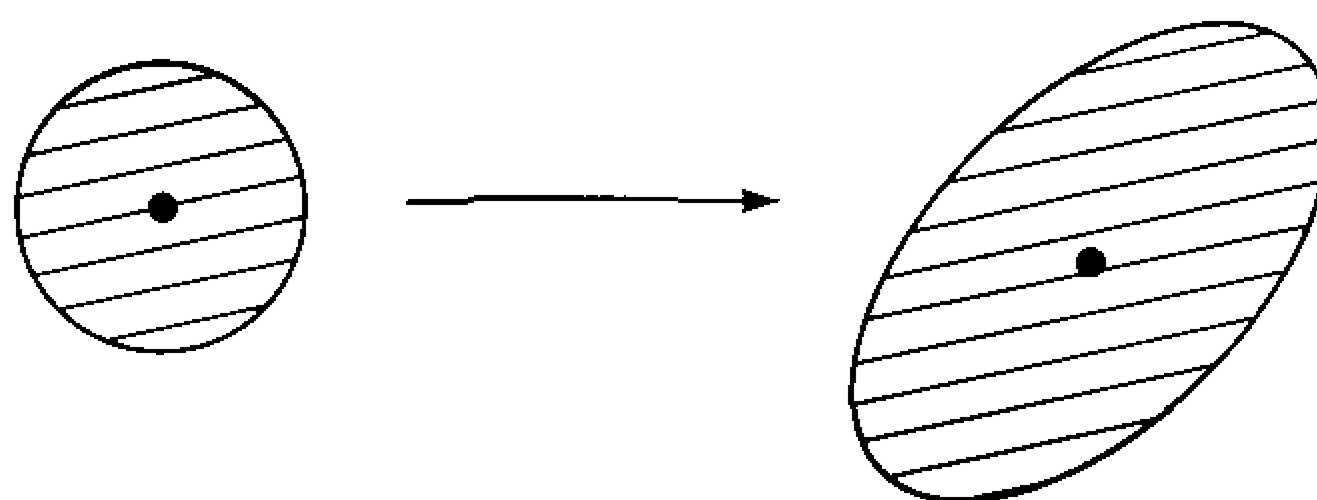
because it has the inverse function $m_{P^{-1}}$. The maps m_P and $m_{P^{-1}}$ are continuous, because matrix multiplication is continuous. Thus m_P is a homeomorphism from G to itself (not a homomorphism). It is also called *left translation* by P, in analogy with translation in the plane, which is left translation in the additive group $\mathbb{R}^{2+}$.

The important property of a group which is implied by the existence of these maps is *homogeneity*. Multiplication by P is a homeomorphism which carries the identity element I to P. So the topological structure of the group G is the same near I as it is near P, and since P is arbitrary, it is the same in the neighborhoods of any two points of the group. This is analogous to the fact that the plane looks the same at any two points.

Left multiplication in SU_2 happens to be defined by an *orthogonal* change of the coordinates (x_1, x_2, x_3, x_4), so it is a rigid motion of the 3-sphere. But multiplication by a matrix needn't be a rigid motion, so the sense in which any group is homogeneous is weaker. For example, let G be the group of real invertible diagonal 2×2 matrices, and let us identify the elements of G with the points (a, d) in the plane, which are not on the coordinate axes. Multiplication by the matrix

$$P = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \tag{7.2}$$

distorts the group G, but it does so continuously.



(7.3) **Figure.** Left multiplication in a group.

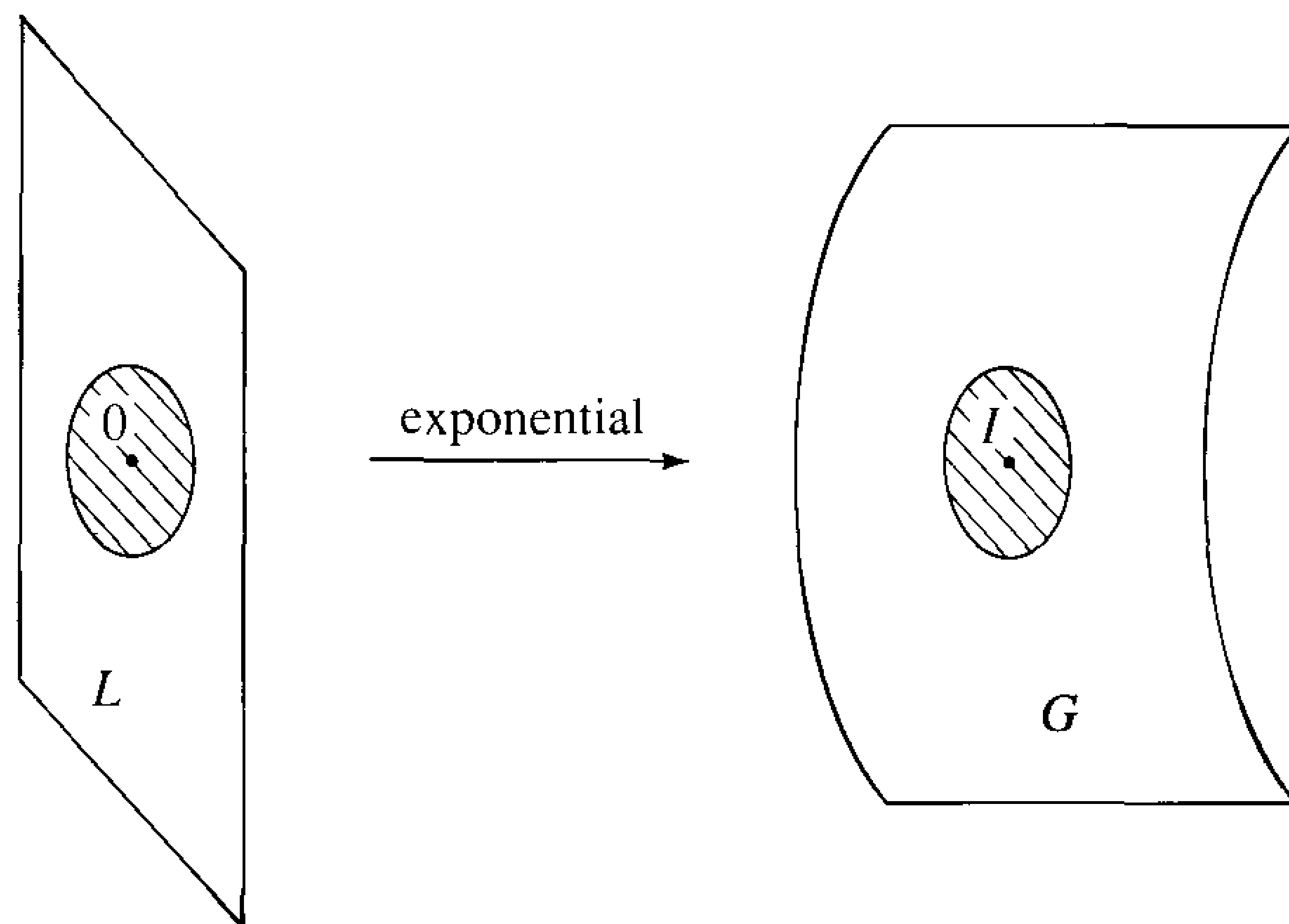
Now the only geometrically reasonable subsets of $\mathbb{R}^k$ which have this homogeneity property are manifolds. A manifold M of dimension d is a subset which is locally homeomorphic to $\mathbb{R}^d$ at any one of its points, meaning that every point $p \in M$ has a neighborhood homeomorphic to an open set in $\mathbb{R}^d$ [see Appendix (3.12)]. It isn't surprising that the classical groups, being homogeneous, are manifolds, though there are subgroups of GL_n which aren't. The group $GL_n(\mathbb{Q})$ of invertible matrices with rational coefficients, for example, is a rather ugly set when viewed geometrically, though it is an interesting group. The following theorem gives a satisfactory answer to the question of which linear groups are manifolds:

(7.4) **Theorem.** Let G be a subgroup of $GL_n(\mathbb{R})$ which is a closed set in $\mathbb{R}^{n\times n}$. Then G is a manifold.

Giving the proof of this theorem here would take us too far afield. Instead, we will illustrate the theorem by showing that the orthogonal groups O_n are manifolds. The proofs for other classical groups are similar.

(7.5) **Proposition.** The orthogonal group O_n is a manifold of dimension $\frac{1}{2}n(n-1)$.

Proof. Let us denote the group O_n by G and denote its Lie algebra, the space of skew-symmetric matrices, by L. Proposition (5.9) tells us that for matrices A near 0, $A \in L$ if and only if $e^A \in G$. Also, the exponential is a homeomorphism from a neighborhood of 0 in $\mathbb{R}^{n\times n}$ to a neighborhood of I. Putting these two facts together, we find that the exponential defines a homeomorphism from a neighborhood of 0 in L to a neighborhood of I in G. Since L is a vector space of dimension $\frac{1}{2}n(n-1)$, it is a manifold. This shows that the condition of being a manifold is satisfied by the orthogonal group at the identity. On the other hand, we saw above that any two points in G have homeomorphic neighborhoods. Therefore G is a manifold, as claimed. □



(7.6) **Figure.**

Here is another application of the principle of homogeneity:

(7.7) **Proposition.** Let G be a path-connected matrix group, and let $H \subset G$ be a subgroup which contains a nonempty open subset of G. Then $H = G$.

Proof. By hypothesis, H contains a nonempty open subset U of G. Since left multiplication by $g \in G$ is a homeomorphism, gU is also open in G. Each translate gU is contained in a single coset of H, namely in gH. Since the translates of U cover G, they cover each coset. In this way, each coset is a union of open subsets of G, and hence it is open itself. So G is partitioned into open subsets—the cosets of H. Now a path-connected set is not a disjoint union of proper open subsets [see Appendix, Proposition (3.11)]. Thus there can be only one coset, and $H = G$. □

We will now apply this proposition to determine the normal subgroups of SU_2.

(7.8) **Theorem.** The only proper normal subgroup of SU_2 is its center $\{\pm I\}$.

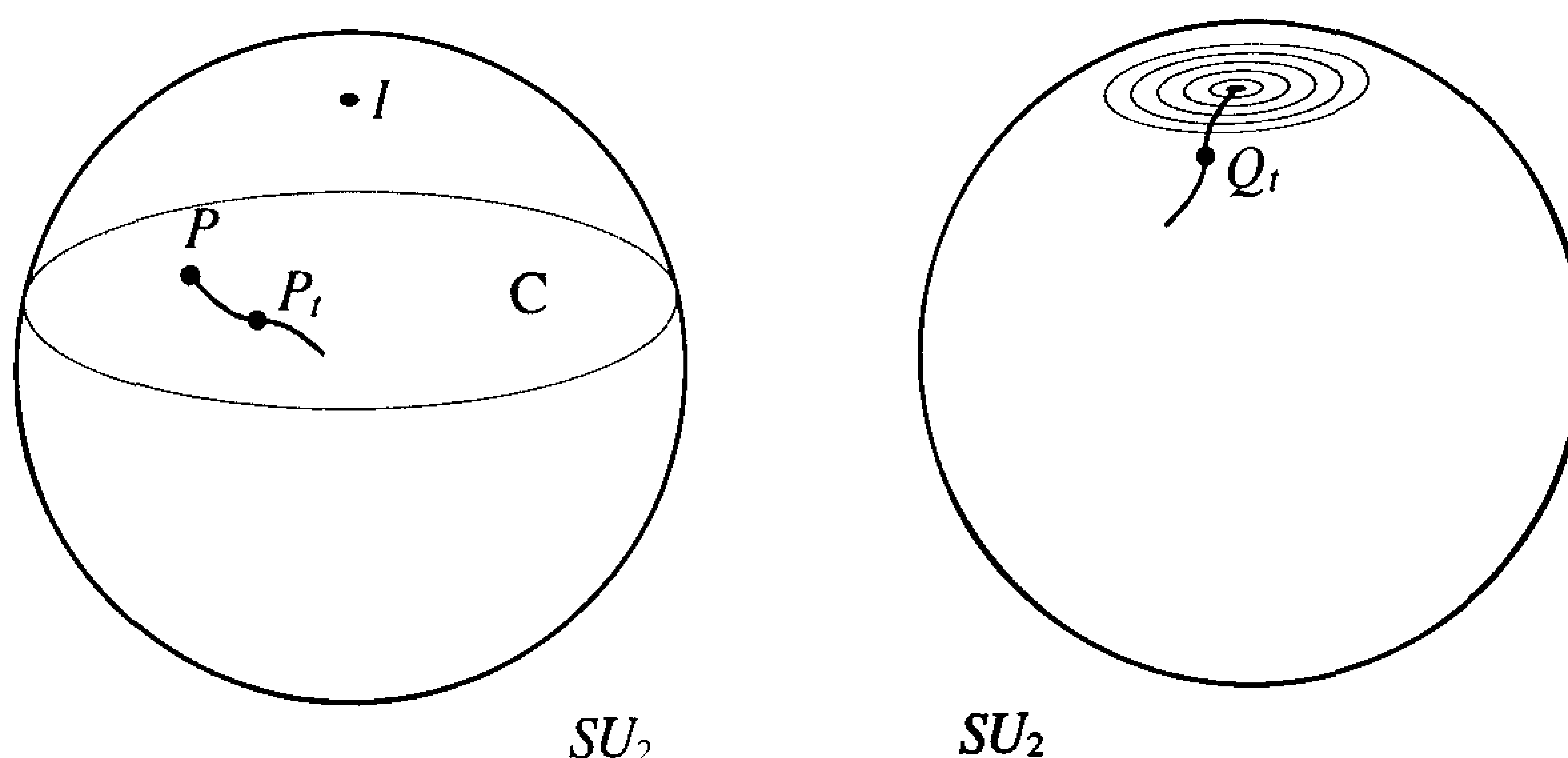
Since there is a surjective map $\varphi: SU_2 \longrightarrow SO_3$ whose kernel is $\{\pm I\}$, the rotation group is isomorphic to a quotient group of SU_2 [Chapter 2 (10.9)]:

$$SO_3 \approx SU_2/\{\pm I\}. \tag{7.9}$$

(7.10) **Corollary.** SO_3 is a simple group; that is, it has no proper normal subgroup.

Proof. The inverse image of a normal subgroup in SO_3 is a normal subgroup of SU_2 which contains $\{\pm I\}$ [Chapter 2 (7.4)]. Theorem (7.8) tells us that there are no proper ones. □

Proof of Theorem (7.8). It is enough to show that if N is a normal subgroup of SU_2 which is not contained in the center $\{\pm I\}$, then N is the whole group. Now since N is normal, it is a union of conjugacy classes [Chapter 6 (2.5)]. And we have



(7.11) **Figure.**

seen that the conjugacy classes are the latitudes, the 2-spheres (2.8). By assumption, N contains a matrix $P \neq \pm I$, so it contains the whole conjugacy class $C = C_P$, which is a 2-sphere. Intuitively, this set looks big enough to generate SU_2. For it has dimension 2 and is not a subgroup. So the set S of all products $P^{-1}Q$ with $P, Q \in C$ is larger than C. Therefore S ought to have dimension 3, which is the dimension of SU_2 itself, so it ought to contain an open set in the group.

To make this intuitive reasoning precise, we choose a nonconstant continuous map from the unit interval $[0, 1]$ to C such that $P_0 = P$ and $P_1 \neq P$. We form the path

$$Q_t = P^{-1}P_t. \tag{7.12}$$

Then $Q_0 = I$, and $Q_1 \neq I$, so this path leads out from I. Since P and P_t are in N, Q_t is in N for every $t \in [0, 1]$. We don't need to know anything else about the path Q_t.

Let $f(t)$ be the function trace Q_t. This is a continuous function on the interval $[0, 1]$. Note that $f(0) = 2$, while $f(1) = \tau < 2$ because $Q_1 \neq I$. By continuity, all values between τ and 2 are taken on by f in the interval.

Since N is normal, it contains the conjugacy class of Q_t for every t. So since trace Q_t takes on all values near 2, Proposition (2.9) tells us that N contains all matrices in SU_2 whose trace is sufficiently near to 2, and this includes all matrices sufficiently near to I. So N contains an open neighborhood of I in SU_2. Now SU_2, being a sphere, is path-connected, so Proposition (7.7) completes the proof. □

We can also apply translation in a group G to tangent vectors. If A is a tangent vector at the identity and if $P \in G$ is arbitrary, then PA is a tangent vector to G at the point P. Intuitively, this is because $P(I + A\epsilon) = P + PA\epsilon$ is the product of elements in G, so it lies in G itself. As always, this heuristic is easy to check for a particular group. We fix A, and associate the tangent vector PA to the element P of G. In this way we obtain what is called a *tangent vector field* on the group G. Since A is nonzero and P is invertible, this vector field does not vanish at any point. Now just the existence of a tangent vector field which is nowhere zero puts strong restrictions on the space G. For example, it is a theorem of topology that any vector field on the 2-sphere must vanish at some point. That is why the 2-sphere has no group structure. But the 3-sphere, being a group, has tangent vector fields which are nowhere zero.

8. SIMPLE GROUPS

Recall that a group G is called *simple* if it is not the trivial group and if it contains no proper normal subgroup (Chapter 6, Section 2). So far, we have seen two nonabelian simple groups: the icosahedral group $I \approx A_5$ [Chapter 6 (2.3)] and the rotation group SO_3 (7.10). This section discusses the classification of simple groups. We will omit most proofs.

Simple groups are important for two reasons. First of all, if a group G has a proper normal subgroup N, then the structure of G is partly described when we

know the structure of N and of the quotient group G/N. If N or G/N has a normal subgroup, we can further decompose the structure of these groups. In this way we may hope to describe a particular finite group G, by building it up inductively from simple groups.

Second, though the condition of being simple is a very strong restriction, simple groups often appear. The classical linear groups are almost simple. For example, we saw in the last section that SU_2 has center $\{\pm I\}$ and that $SU_2/\{\pm I\} \approx SO_3$ is a simple group. The other classical groups have similar properties.

In order to focus attention, we will restrict our discussion here to the complex groups. We will use the symbol Z to denote the center of any group. The following theorem would take too much time to prove here, but we will illustrate it in the special case of $SL_2(\mathbb{C})$.

(8.1) **Theorem.**

(a) The center Z of the special linear group $SL_n(\mathbb{C})$ is a cyclic group, generated by the matrix ζI where $\zeta = e^{2\pi i/n}$. The quotient group $SL_n(\mathbb{C})/Z$ is simple if $n \geq 2$.

(b) The center Z of the complex special orthogonal group $SO_n(\mathbb{C})$ is $\{\pm I\}$ if n is even, and is the trivial group $\{I\}$ if n is odd. The group SO_n/Z is simple if $n = 3$ or if $n \geq 5$.

(c) The center Z of the symplectic group $SP_{2n}(\mathbb{C})$ is $\{\pm I\}$, and $SP_{2n}(\mathbb{C})/Z$ is simple if $n \geq 1$. □

The group $SL_n(\mathbb{C})/Z$ is called the *projective group* and is denoted by $PSL_n(\mathbb{C})$:

(8.2) $$PSL_n(\mathbb{C}) = SL_n(\mathbb{C})/Z, \quad \text{where } Z = \{\zeta I \mid \zeta^n = 1\}.$$

To illustrate Theorem (8.1), we will prove that $PSL_2(\mathbb{C}) = SL_2(\mathbb{C})/\{\pm I\}$ is simple. In fact, we will show that $PSL_2(F)$ is a simple group for almost all fields F.

(8.3) **Theorem.** Let F be a field which is not of characteristic 2 and which contains at least seven elements. Then the only proper normal subgroup of $SL_2(F)$ is the subgroup $\{\pm I\}$. Thus $PSL_2(F) = SL_2(F)/\{\pm I\}$ is a simple group.

Since the center of $SL_2(F)$ is a normal subgroup, it follows from the theorem that it is the group $\{\pm I\}$.

(8.4) **Corollary.** There are infinitely many nonabelian finite simple groups.

Proof of Theorem (8.3). The proof is algebraic, but it is closely related to the geometric proof given for the analogous assertion for SU_2 in the last section. Our procedure is to conjugate and multiply until the group is generated. To simplify notation, we will denote $SL_2(F)$ by SL_2. Let N be a normal subgroup of SL_2 which contains a matrix $A \neq \pm I$. We must show that $N = SL_2$. Since one possibility is that N

is the normal subgroup generated by A and its conjugates, we must show that the conjugates of this one matrix suffice to generate the whole group.

The first step in our proof will be to show that N contains a triangular matrix different from $\pm I$. Now if our given matrix A has eigenvalues in the field F, then it will be conjugate to a triangular matrix. But since we want to handle arbitrary fields, we can not make this step so easily. Though easy for the complex numbers, this step is the hardest part of the proof for a general field.

(8.5) **Lemma.** *N contains a triangular matrix $A \neq \pm I$.*

Proof. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a matrix in N which is different from $\pm I$. If $c = 0$, then A is the required matrix.

Suppose that $c \neq 0$. In this case, we will construct a triangular matrix out of A and its conjugates. We first compute the conjugate

$$\begin{bmatrix} 1 & x \\ & 1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & -x \\ & 1 \end{bmatrix} = \begin{bmatrix} a+xc & * \\ c & d-xc \end{bmatrix} = A'.$$

Since $c \neq 0$, we may choose x so that $a + xc = 0$. The matrix A' is in N, so N contains a matrix whose upper left entry is zero. We replace A by this matrix, so that it has the form $A = \begin{bmatrix} & b \\ c & d \end{bmatrix}$. Unfortunately the zero is in the wrong place.

Note that since $\det A = 1$, $bc = -1$ in our new matrix A. We now compute the commutator $P^{-1}A^{-1}PA$ with a diagonal matrix:

$$P^{-1}A^{-1}PA = \begin{bmatrix} u & \\ & u^{-1} \end{bmatrix}\begin{bmatrix} d & -b \\ -c & \end{bmatrix}\begin{bmatrix} u^{-1} & \\ & u \end{bmatrix}\begin{bmatrix} & b \\ c & d \end{bmatrix} = \begin{bmatrix} u^2 & (1-u^2)bd \\ & u^{-2} \end{bmatrix}.$$

This matrix, which is in our normal subgroup N, is as required unless it is $\pm I$. If so, then $u^2 = \pm 1$ and $u^4 = 1$. But we are free to form the matrix P with an arbitrary element u in $F^{\times}$. We will show [Chapter 11 (1.8)] that the polynomial $x^4 - 1$ has at most four roots in any field. So there are at most four elements $u \in F$ with $u^4 - 1$. Our hypothesis is that $F^{\times}$ contains at least five elements. So we can choose $u \in F^{\times}$ with $u^4 \neq 1$. Then $P^{-1} A^{-1} PA$ is the required matrix. □

(8.6) **Lemma.** *N contains a matrix of the form $\begin{bmatrix} 1 & u \\ & 1 \end{bmatrix}$, with $u \neq 0$.*

Proof. By the previous lemma, N contains a triangular matrix $A = \begin{bmatrix} a & b \\ & d \end{bmatrix} \neq \pm I$. If $d \neq a$, let $A' = \begin{bmatrix} a & b' \\ & d \end{bmatrix}$ be its conjugate by the matrix $\begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}$. Then $b' = b + d - a$. Since $\det A = ad = 1$, the product

$$A'^{-1}A = \begin{bmatrix} d & -b' \\ & a \end{bmatrix}\begin{bmatrix} a & b \\ & d \end{bmatrix} = \begin{bmatrix} 1 & ad-d^2 \\ & 1 \end{bmatrix}$$

is the required matrix. If $a = d$, then $a = \pm 1$ because $\det A = 1$, and it follows that $b \neq 0$. In this case, one of the two matrices A or A^2 is as required. □

(8.7) **Lemma.** Let F be a field. The conjugacy class in SL_2 of the matrix $\begin{bmatrix} 1 & u \\ & 1 \end{bmatrix}$ contains the matrices $\begin{bmatrix} 1 & \\ -u & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & a^2u \\ & 1 \end{bmatrix}$, for all $a \neq 0$.

Proof.

$$\begin{bmatrix} & -1 \\ 1 & \end{bmatrix}\begin{bmatrix} 1 & u \\ & 1 \end{bmatrix}\begin{bmatrix} & 1 \\ -1 & \end{bmatrix} = \begin{bmatrix} 1 & \\ -u & 1 \end{bmatrix} \text{ and } \begin{bmatrix} a & \\ & a^{-1} \end{bmatrix}\begin{bmatrix} 1 & u \\ & 1 \end{bmatrix}\begin{bmatrix} a^{-1} & \\ & a \end{bmatrix} = \begin{bmatrix} 1 & a^2u \\ & 1 \end{bmatrix}. \square$$

(8.8) **Lemma.** Let F be a field of characteristic $\neq 2$. The additive group F^+ of the field is generated by the squares of elements of F.

Proof. We show that every element $x \in F$ can be written in the form $a^2 - b^2 = (a + b)(a - b)$, with $a, b \in F$. To do this, we solve the system of linear equations $a + b = 1$, $a - b = x$. This is where the assumption that the characteristic of F is not 2 is used. In characteristic 2, these equations need not have a solution. □

(8.9) **Lemma.** Let F be a field of characteristic $\neq 2$. If a normal subgroup N of $SL_2(F)$ contains a matrix $\begin{bmatrix} 1 & u \\ & 1 \end{bmatrix}$ with $u \neq 0$, then it contains all such matrices.

Proof. The set of x such that $\begin{bmatrix} 1 & x \\ & 1 \end{bmatrix} \in N$ is a subgroup of F^+, call it S. We want to show that $S = F^+$. Lemma (8.7) shows that if $u \in S$, then $a^2u \in S$ for all $a \in F$. Since the squares generate F^+, the set of elements $\{a^2u \mid a \in F\}$ generates the additive subgroup F^+u of F^+, and this subgroup is equal to F^+ because u is invertible. Thus $S = F^+$, as required. □

(8.10) **Lemma.** For every field F, the group $SL_2(F)$ is generated by the elementary matrices $\begin{bmatrix} 1 & u \\ & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & \\ u & 1 \end{bmatrix}$.

Proof. We perform row reduction on a matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(F)$, using only the matrices of this form. We start work on the first column, reducing it to e_1. We eliminate the case $c = 0$ by adding the first row to the second if necessary. Then we add a multiple of the second row to the first to change a to 1. Finally, we clear out the entry c. At this point, the matrix has the form $A' = \begin{bmatrix} 1 & b' \\ 0 & d' \end{bmatrix}$. Then $d' = \det A' = \det A = 1$, and we can clear out the entry b', ending up with the identity matrix. Since we needed four operations or less to reduce to the identity, A is a product of at most four of these elementary matrices. □

The proof of Theorem (8.3) is completed by combining Lemmas (8.6), (8.7), (8.9), and (8.10). □

A famous theorem of Cartan asserts that the list (8.1) of simple groups is almost complete. Of course there are other simple groups; for instance, we have just proved that $PSL_2(F)$ is simple for most fields F. But if we restrict ourselves to complex algebraic groups, the list of simple groups becomes very short.

A subgroup G of $GL_n(\mathbb{C})$ is called a *complex algebraic group* if it is the set of solutions of a finite system of polynomial equations in the matrix entries. This is analogous to the concept of a real algebraic group introduced in Section 6. It will not be apparent why the property of being defined by polynomial equations is a reasonable one, but one thing is easy to see: Except for the unitary groups U_n and SU_n, all the complex classical groups are complex algebraic groups.

(8.11) **Theorem.**

(a) The groups $PSL_n(\mathbb{C}) = SL_n(\mathbb{C})/Z$, $SO_n(\mathbb{C})/Z$, and $SP_{2n}(\mathbb{C})/Z$ are path-connected complex algebraic groups.

(b) In addition to the isomorphism classes of these groups, there are exactly five isomorphism classes of simple, path-connected complex algebraic groups, called the *exceptional groups*.

Theorem (8.11) is too hard to prove here. It is based on a classification of the corresponding Lie algebras. What we should learn is that there are not many simple algebraic groups. This ought to be reassuring after the last chapter, where structures on a vector space were introduced one after the other, each with its own group of symmetries. There seemed to be no end. Now we see that we actually ran across most of the possible symmetry types, at least those associated to *simple* algebraic groups. It is no accident that these structures are important. □

A large project, the classification of the *finite* simple groups, was completed in 1980. The finite simple groups we have seen are the groups of prime order, the icosahedral group $I \approx A_5$ [Chapter 6 (2.3)], and the groups $PSL_2(F)$ where F is a finite field (8.3), but there are many more. The alternating groups A_n are simple for all $n \geq 5$.

Linear groups play a dominant role in the classification of the finite simple groups as well as of the complex algebraic groups. Each of the forms (8.11) leads to a whole series of finite simple groups when finite fields are substituted for the complex field. Also, some finite simple groups are analogous to the unitary groups. All of these finite linear groups are said to be of *Lie type*.

According to Theorem (8.3), $PSL_2(\mathbb{F}_7)$ is a finite simple group; its order is 168. This is the second smallest simple group; A_5 is the smallest. The orders of the smallest nonabelian simple groups are

$$60, 168, 360, 504, 660, 1092, 2448. \tag{8.12}$$

For each of these seven integers N, there is a single isomorphism class of simple groups of order N, and it is represented by $PSL_2(F)$ for a suitable finite field F. [The alternating group A_5 happens to be isomorphic to $PSL_2(\mathbb{F}_5)$.]

In addition to the groups of prime order, the alternating groups, and the groups of Lie type, there are exactly 26 finite simple groups called the *sporadic groups*. The smallest sporadic group is the *Matthieu group* M_{11}, whose order is 7920. The largest is called the *Monster*; its order is roughly 10^{53}. So the finite simple groups form a list which, though longer, is somewhat analogous to the list (8.11) of simple algebraic groups.

It seems unfair to crow about the successes of a theory and to sweep all its failures under the rug.

Richard Brauer

EXERCISES

1. The Classical Linear Groups

1. (a) Find a subgroup of $GL_2(\mathbb{R})$ which is isomorphic to $\mathbb{C}^\times$.
   (b) Prove that for every n, $GL_n(\mathbb{C})$ is isomorphic to a subgroup of $GL_{2n}(\mathbb{R})$.
2. Show that $SO_2(\mathbb{C})$ is not a bounded set in $\mathbb{C}^4$.
3. Prove that $SP_2(\mathbb{R}) = SL_2(\mathbb{R})$, but that $SP_4(\mathbb{R}) \neq SL_4(\mathbb{R})$.
4. According to Sylvester's Law, every 2×2 real symmetric matrix is congruent to exactly one of six standard types. List them. If we consider the operation of $GL_2(\mathbb{R})$ on 2×2 matrices by $P, A \rightsquigarrow PAP^t$, then Sylvester's Law asserts that the symmetric matrices form six orbits. We may view the symmetric matrices as points in $\mathbb{R}^3$, letting (x,y,z) correspond to the matrix $\begin{bmatrix} x & y \\ y & z \end{bmatrix}$. Find the decomposition of $\mathbb{R}^3$ into orbits explicitly, and make a clear drawing showing the resulting geometric configuration.
5. A matrix P is orthogonal if and only if its columns form an orthonormal basis. Describe the properties that the columns of a matrix must have in order for it to be in the Lorentz group $O_{3,1}$.
6. Prove that there is no continuous isomorphism from the orthogonal group O_4 to the Lorentz group $O_{3,1}$.
7. Describe by equations the group $O_{1,1}$, and show that it has four connected components.
8. Describe the orbits for the operation of $SL_2(\mathbb{R})$ on the space of real symmetric matrices by $P, A \rightsquigarrow PAP^t$.
9. Let F be a field whose characteristic is not 2. Describe the orbits for the action $P, A \rightsquigarrow PAP^t$ of $GL_2(F)$ on the space of 2×2 symmetric matrices with coefficients in F.
10. Let $F = \mathbb{F}_2$. Classify the orbits of $GL_n(F)$ for the action on the space of symmetric $n \times n$ matrices by finding representatives for each congruence class.

11. Prove that the following matrices are symplectic, if the blocks are $n \times n$:
$\begin{bmatrix} & -I \\ I & \end{bmatrix}$, $\begin{bmatrix} A^t & \\ & A^{-1} \end{bmatrix}$, $\begin{bmatrix} I & B \\ & I \end{bmatrix}$, where $B = B^t$ and A is invertible.

12. Prove that the symplectic group $SP_{2n}(\mathbb{R})$ operates transitively on $\mathbb{R}^{2n}$.

*13. Prove that $SP_{2n}(\mathbb{R})$ is path-connected, and conclude that every symplectic matrix has determinant 1.

2. The Special Unitary Group SU_2

1. Let P, Q be elements of SU_2, represented by the real vectors (x_1, x_2, x_3, x_4), (y_1, y_2, y_3, y_4). Compute the real vector which corresponds to the product PQ.
2. Prove that the subgroup SO_2 of SU_2 is conjugate to the subgroup T of diagonal matrices.
3. Prove that SU_2 is path-connected. Do the same for SO_3.
4. Prove that U_2 is homeomorphic to the product $S^3 \times S^1$.
5. Let G be the group of matrices of the form $\begin{bmatrix} x & y \\ & 1 \end{bmatrix}$, where $x, y \in \mathbb{R}$ and $x > 0$. Determine the conjugacy classes in G, and draw them in the (x, y)-plane.

*6. (a) Prove that every element P (2.4) of SU_2 can be written as a product: $P = DRD'$, where $D, D' \in T$ (2.13), and $R \in SO_2$ is a rotation through an angle θ with $0 \leq \theta \leq \pi/2$.
   (b) Assume that the matrix entries a, b of P are not zero. Prove that this representation is unique, except that the pair D, D' can be replaced by $-D, -D'$.
   (c) Describe the double cosets TPT, $P \in SU_2$. Prove that if the entries a, b of P are not zero, then the double coset is homeomorphic to a torus, and describe the remaining double cosets.

3. The Orthogonal Representation of SU_2

1. Compute the stabilizer H of the matrix $\begin{bmatrix} & 1 \\ 1 & \end{bmatrix}$ for the action of conjugation by SU_2, and describe $\varphi(P)$ for $P \in H$.
2. Prove that every great circle in SU_2 is a coset of one of the longitudes (2.14).
3. Find a subset of $\mathbb{R}^3$ which is homeomorphic to the space $S \times \Theta$ of (3.4).
4. Derive a formula for $\langle A, A \rangle$ in terms of the determinant of A.
5. The rotation group SO_3 may be mapped to the 2-sphere by sending a rotation matrix to its first column. Describe the fibres of this map.
6. Extend the map φ defined in this section to a homomorphism $\Phi: U_2 \longrightarrow SO_3$, and describe the kernel of Φ.
7. Prove by direct computation that the matrix (3.11) is in SO_3.

*8. Describe the conjugacy classes in SO_3 carefully, relating them to the conjugacy classes of SU_2.

9. Prove that the operation of SU_2 on any conjugacy class other than $\{I\}$, $\{-I\}$ is by rotations of the sphere.
10. Find a bijective correspondence between elements of SO_3 and pairs (p, v) consisting of a point p on the unit 2-sphere S and a unit tangent vector v to S at p.

11. Prove Proposition (3.20).

*12. (a) Calculate left multiplication by a fixed matrix P in SU_2 explicitly in terms of the coordinates x_1, x_2, x_3, x_4. Prove that it is multiplication by a 4×4 orthogonal matrix Q, hence that it is a rigid motion of the unit 3-sphere S^3.

(b) Prove that Q is orthogonal by a method similar to that used in describing the orthogonal representation: Express dot product of the vectors (x_1, x_2, x_3, x_4), (x_1', x_2', x_3', x_4') corresponding to two matrices $P, P' \in SU_2$ in terms of matrix operations.

(c) Determine the matrix which describes the operation of conjugation by a fixed matrix P on SU_2.

*13. (a) Let H_i be the subgroup of SO_3 of rotations about the x_i-axis, $i = 1, 2, 3$. Prove that every element of SO_3 can be written as a product ABA', where $A, A' \in H_1$ and $B \in H_2$. Prove that this representation is unique unless $B = I$.

(b) Describe the double cosets $H_1 Q H_1$ geometrically.

*14. Let H_i be the subgroup of SO_3 of rotations about the x_i-axis. Prove that every element $Q \in SO_3$ can be written in the form $A_1 A_2 A_3$, with $A_i \in H_i$.

4. *The Special Linear Group* $SL_2(\mathbb{R})$

1. Let $G = SL_2(\mathbb{C})$. Use the operation of G on rays $\{rX\} \mid r \in \mathbb{R}, r > 0\}$ in $\mathbb{C}^2$ to prove that G is homeomorphic to the product $SU_2 \times H$, where H is the stabilizer of the ray $\{re_1\}$, and describe H explicitly.

2. (a) Prove that the rule $P, A \rightsquigarrow PAP^*$ defines an operation of $SL_2(\mathbb{C})$ on the space W of all hermitian matrices.

   (b) Prove that the function $\langle A, A' \rangle = \det(A + A') - \det A - \det A'$ is a bilinear form on W, whose signature is $(3, 1)$.

   (c) Use (a) and (b) to define a homomorphism $\varphi: SL_2(\mathbb{C}) \longrightarrow O_{3,1}$, whose kernel is $\{\pm I\}$.

   *(d) Prove that the image of φ is the connected component of the identity in $O_{3,1}$.

3. Let P be a matrix in $SO_3(\mathbb{C})$.

   (a) Prove that 1 is an eigenvalue of P.

   (b) Let X_1, X_2 be eigenvectors for P, with eigenvalues λ_1, λ_2. Prove that $X_1{}^t X_2 = 0$, unless $\lambda_1 = \lambda_2^{-1}$.

   (c) Prove that if X is an eigenvector with eigenvalue 1 and if $P \neq I$, then $X^t X \neq 0$.

4. Let $G = SO_3(\mathbb{C})$.

   (a) Prove that left multiplication by G is a transitive operation on the set of vectors X such that $X^t X = 1$.

   (b) Determine the stabilizer of e_1 for left multiplication by G.

   (c) Prove that G is path-connected.

5. *One-Parameter Subgroups*

1. Determine the differentiable homomorphisms from $\mathbb{C}^+$ to $SL_n(\mathbb{C})$.
2. Describe all one-parameter subgroups of $\mathbb{C}^\times$.
3. Describe by equations the images of all one-parameter subgroups of the group of real 2×2 diagonal matrices, and make a neat drawing showing them.
4. Let $\varphi: \mathbb{R}^+ \longrightarrow GL_n(\mathbb{R})$ be a one-parameter subgroup. Prove that $\ker \varphi$ is either trivial, or the whole group, or else it is infinite cyclic.

5. Find the conditions on a matrix A so that e^{tA} is a one-parameter subgroup of the special unitary group SU_n, and compute the dimension of that group.

6. Let G be the group of real matrices of the form $\begin{bmatrix} x & y \\ & 1 \end{bmatrix}$, $x > 0$.
   (a) Determine the matrices A such that e^{tA} is a one-parameter subgroup of G.
   (b) Compute e^A explicitly for the matrices determined in (a).
   (c) Make a drawing showing the one-parameter subgroups in the (x, y)-plane.

7. Prove that the images of the one-parameter subgroups of SU_2 are the conjugates of T (see Section 3). Use this to give an alternative proof of the fact that these conjugates are the longitudes.

8. Determine the one-parameter subgroups of U_2.

9. Let $\varphi(t) = e^{tA}$ be a one-parameter subgroup of G. Prove that the cosets of im φ are matrix solutions of the differential equation $dX/dt = AX$.

10. Can a one-parameter subgroup of $GL_n(\mathbb{R})$ cross itself?

*11. Determine the differentiable homomorphisms from SO_2 to $GL_n(\mathbb{R})$.

6. The Lie Algebra

1. Compute $(A + B\epsilon)^{-1}$, assuming that A is invertible.

2. Compute the infinitesimal tangent vectors to the plane curve $y^2 = x^3$ at the point (1,1) and at the point (0,0).

3. (a) Sketch the curve C: $x_2^2 = x_1^3 - x_1^2$.
   (b) Prove that this locus is a manifold of dimension 1 if the origin is deleted.
   (c) Determine the tangent vectors and the infinitesimal tangents to C at the origin.

4. Let S be a real algebraic set defined by one equation $f = 0$.
   (a) Show that the equation $f^2 = 0$ defines the same locus S.
   (b) Show that $\nabla(f^2)$ vanishes at every point x of S, hence that every vector is an infinitesimal tangent at x, when the defining equation is taken to be $f^2 = 0$.

5. Show that the set defined by $xy = 1$ is a subgroup of the group of diagonal matrices $\begin{bmatrix} x & \\ & y \end{bmatrix}$, and compute its Lie algebra.

6. Determine the Lie algebra of the unitary group.

7. (a) Prove the formula $\det(I + A\epsilon) = 1 + \operatorname{trace} A\epsilon$.
   (b) Let A be an invertible matrix. Compute $\det(A + B\epsilon)$.

8. (a) Show that O_2 operates by conjugation on its Lie algebra.
   (b) Show that the operation in (a) is compatible with the bilinear form $\langle A, B\rangle = \frac{1}{2}\operatorname{trace} AB$.
   (c) Use the operation in (a) to define a homomorphism $O_2 \longrightarrow O_2$, and describe this homomorphism explicitly.

9. Compute the Lie algebra of the following: (a) U_n; (b) SU_n; (c) $O_{3,1}$; (d) $SO_n(\mathbb{C})$. In each case, show that e^{tA} is a one-parameter subgroup if and only if $I + A\epsilon$ lies in the group.

*10. Determine the Lie algebra of $G = SP_{2n}(\mathbb{R})$, using block form $M = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$.

11. (a) Show that $\mathbb{R}^3$ becomes a Lie algebra if the bracket is defined to be the cross product $[X, Y] = X \times Y = (x_2y_3 - y_2x_3, x_3y_1 - y_1x_3, x_1y_2 - x_2y_1)$.
    (b) Show that this Lie algebra is isomorphic to the Lie algebra of SO_3.

12. Classify all complex Lie algebras of dimension ≤ 3.

*13. The *adjoint representation* of a linear group G is the representation by conjugation on its Lie algebra: $G \times L \longrightarrow L$ is defined to be $P, A \rightsquigarrow PAP^{-1}$. The form $\langle A, A' \rangle = \text{trace}(AA')$ on L is called the *Killing form*. For each of the following groups, verify that if $P \in G$ and $A \in L$, then $PAP^{-1} \in L$, and prove that the Killing form is symmetric and bilinear and that the operation is compatible with the form, i.e., that $\langle A, A' \rangle = \langle PAP^{-1}, PA'P^{-1} \rangle$.
(a) SO_n (b) SU_n (c) $O_{3,1}$ (d) $SO_n(\mathbb{C})$ (e) $SP_{2n}(\mathbb{R})$

14. Prove that the Killing form is negative definite on the Lie algebra of (a) SU_n and (b) SO_n.

15. Determine the signature of the Killing form on the Lie algebra of $SL_n(\mathbb{R})$.

16. (a) Use the adjoint representation of SU_n to define a homomorphism $\varphi: SU_n \longrightarrow SO_m$, where $m = n^2 - 1$.
(b) Show that when $n = 2$, this representation is equivalent to the orthogonal representation defined in Section 3.

17. Use the adjoint representation of $SL_2(\mathbb{C})$ to define an isomorphism $SL_2(\mathbb{C})/\{\pm I\} \approx SO_3(\mathbb{C})$.

7. Translation in a Group

1. Compute the dimensions of the following groups.
(a) SU_n (b) $SO_n(\mathbb{C})$ (c) $SP_{2n}(\mathbb{R})$ (d) $O_{3,1}$

2. Using the exponential, find all solutions near I of the equation $P^2 = I$.

3. Find a path-connected, nonabelian subgroup of $GL_2(\mathbb{R})$ of dimension 2.

4. (a) Show that every positive definite hermitian matrix A is the square of another positive definite hermitian matrix B.
(b) Show that B is uniquely determined by A.

*5. Let A be a nonsingular matrix, and let B be a positive definite hermitian matrix such that $B^2 = AA^*$.
(a) Show that A^*B^{-1} is unitary.
(b) Prove the *Polar decomposition:* Every nonsingular matrix A is a product $A = PU$, where P is positive definite hermitian and U is unitary.
(c) Prove that the Polar decomposition is unique.
(d) What does this say about the operation of left multiplication by the unitary group U_n on the group GL_n?

*6. State and prove an analogue of the Polar decomposition for real matrices.

*7. (a) Prove that the exponential map defines a bijection between the set of all hermitian matrices and the set of positive definite hermitian matrices.
(b) Describe the topological structure of $GL_2(\mathbb{C})$ using the Polar decomposition and (a).

8. Let B be an invertible matrix. Describe the matrices A such that $P = e^A$ is in the centralizer of B.

*9. Let S denote the set of matrices $P \in SL_2(\mathbb{R})$ with trace r. These matrices can be written in the form $\begin{bmatrix} x & y \\ z & r-x \end{bmatrix}$, where (x, y, z) lies on the quadric $x(r - x) - yz = 1$.
(a) Show that the quadric is either a hyperbola of one or two sheets, or else a cone, and determine the values of r which correspond to each type.
(b) In each case, determine the decomposition of the quadric into conjugacy classes.

(c) Extend the method of proof of Theorem (7.11) to show that the only proper normal subgroup of $SL_2(\mathbb{R})$ is $\{\pm I\}$.

10. Draw the tangent vector field PA to the group $\mathbb{C}^\times$, when $A = 1 + i$.

8. Simple Groups

1. Which of the following subgroups of $GL_n(\mathbb{C})$ are complex algebraic groups?
(a) $GL_n(\mathbb{Z})$ (b) SU_n (c) upper trianguiar matrices

2. (a) Write the polynomial functions in the matrix entries which define $SO_n(\mathbb{C})$.
(b) Write out the polynomial equations which define the symplectic group.
(c) Show that the unitary group U_n can be defined by real polynomial equations in the real and imaginary parts of the matrix entries.

3. Determine the centers of the groups $SL_n(\mathbb{R})$ and $SL_n(\mathbb{C})$.

4. Describe isomorphisms (a) $PSL_2(\mathbb{F}_2) \approx S_3$ and (b) $PSL_2(\mathbb{F}_3) \approx A_4$.

5. Determine the conjugacy classes of $GL_2(\mathbb{F}_3)$.

6. Prove that $SL_2(F) = PSL_2(F)$ for any field F of characteristic 2.

7. (a) Determine all normal subgroups of $GL_2(\mathbb{C})$ which contain its center $Z = \{cI\}$.
(b) Do the same for $GL_2(\mathbb{R})$.

8. For each of the seven orders (8.12), determine the order of the field F such that $PSL_2(F)$ has order n.

***9.** Prove that there is a simple group of order 3420.

10. (a) Let Z be the center of $GL_n(\mathbb{C})$. Is $PSL_n(\mathbb{C})$ isomorphic to $GL_n(\mathbb{C})/Z$?
(b) Answer the same question as in (a), with $\mathbb{R}$ replacing $\mathbb{C}$.

11. Prove that $PSL_2(\mathbb{F}_5)$ is isomorphic to A_5.

***12.** Analyze the proof of Theorem (8.3) to prove that $PSL_2(F)$ is a simple group when F is a field of characteristic 2, except for the one case $F = \mathbb{F}_2$.

13. (a) Let P be a matrix in the center of SO_n, and let A be a skew-symmetric matrix. Prove that $PA = AP$ by differentiating the matrix function e^{At}.
(b) Prove that the center of SO_n is trivial if n is odd and is $\{\pm I\}$ if n is even and ≥ 4.

14. Compute the orders of the following groups.
(a) $SO_2(\mathbb{F}_3)$ and $SO_3(\mathbb{F}_3)$
(b) $SO_2(\mathbb{F}_5)$ and $SO_3(\mathbb{F}_5)$

***15.** (a) Consider the operation of $SL_2(\mathbb{C})$ by conjugation on the space V of complex 2×2 matrices. Show that with the basis $(e_{11}, e_{12}, e_{21}, e_{22})$ of V, the matrix of conjugation by $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has the block form $\begin{bmatrix} aB & bB \\ cB & dB \end{bmatrix}$, where $B = (A^t)^{-1} = \begin{bmatrix} d & -c \\ -b & a \end{bmatrix}$.
(b) Prove that this operation defines a homomorphism $\varphi: SL_2(\mathbb{C}) \longrightarrow GL_4(\mathbb{C})$, and that the image of φ is isomorphic to $PSL_2(\mathbb{C})$.
(c) Prove that $PSL_2(\mathbb{C})$ is an algebraic group by finding polynomial equations in the entries y_{ij} of a 4×4 matrix whose solutions are precisely the matrices in im φ.

***16.** Prove that $PSL_n(\mathbb{C})$ is a simple group.

***17.** There is no simple group of order $2^5 \cdot 7 \cdot 11$. Assuming this, determine the next smallest order after 2448 for a nonabelian simple group.

Miscellaneous Exercises

1. *Quaternions* are expressions of the form $\alpha = a + bi + cj + dk$, where $a, b, c, d \in \mathbb{R}$. They can be added and multiplied using the rules of multiplication for the quaternion group [Chapter 2 (2.12)].

(a) Let $\overline{\alpha} = a - bi - cj - dk$. Compute $\alpha\overline{\alpha}$.

(b) Prove that every $\alpha \neq 0$ has a multiplicative inverse.

(c) Prove that the set of quaternions α such that $a^2 + b^2 + c^2 + d^2 = 1$ forms a group under multiplication which is isomorphic to SU_2.

2. The *affine group* $A_n = A_n(\mathbb{R})$ is the group of coordinate changes in $(x_1, \ldots, x_n)$ which is generated by $GL_n(\mathbb{R})$ and by the group T_n of translations: $t_a(x) = x + a$. Prove that T_n is a normal subgroup of A_n and that A_n/T_n is isomorphic to $GL_n(\mathbb{R})$.

3. *Cayley Transform*: Let U denote the set of matrices A such that $I + A$ is invertible, and define $A' = (I - A)(I + A)^{-1}$.

(a) Prove that if $A \in U$, then $A' \in U$, and prove that $A'' = A$.

(b) Let V denote the vector space of real skew-symmetric $n \times n$ matrices. Prove that the rule $A \rightsquigarrow (I - A)(I + A)^{-1}$ defines a homeomorphism from a neighborhood of 0 in V to a neighborhood of I in SO_n.

(c) Find an analogous statement for the unitary group.

(d) Let $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$. Show that a matrix $A \in U$ is symplectic if and only if $A'^t J = -JA'$.

***4.** Let $p(t) = t^2 - ut + 1$ be a quadratic polynomial, with coefficients in the field $F = \mathbb{F}_p$.

(a) Prove that if p has two distinct roots in F, then the matrices with characteristic polynomial p form two conjugacy classes in $SL_2(F)$, and determine their orders.

(b) Prove that if p has two equal roots, then the matrices with characteristic polynomial p form three conjugacy classes in $SL_n(F)$, and determine their orders.

(c) Suppose that p has no roots in F. Determine the centralizer of the matrix $A = \begin{bmatrix} & -1 \\ 1 & u \end{bmatrix}$ in $SL_2(F)$, and compute the order of the conjugacy class of A.

(d) Find the class equations of $SL_2(\mathbb{F}_3)$ and $SL_2(\mathbb{F}_5)$.

(e) Find the class equations of $PSL_2(\mathbb{F}_3)$ and $PSL_2(\mathbb{F}_5)$, and reconcile your answer with the class equations of A_4 and A_5.

(f) Compute the class equation for $SL_2(\mathbb{F}_7)$ and for $PSL_2(\mathbb{F}_7)$. Use the class equation for $PSL_2(\mathbb{F}_7)$ to show that this group is simple.

Chapter 9

Group Representations

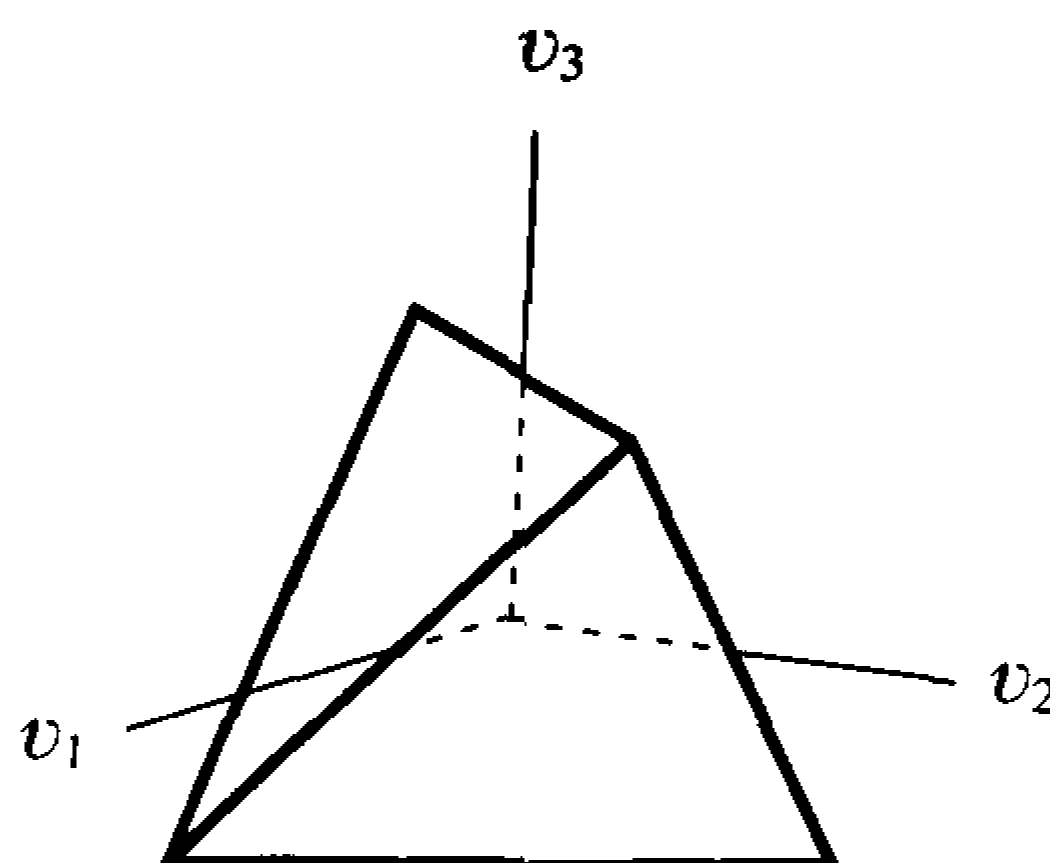
A tremendous effort has been made by mathematicians
for more than a century to clear up the chaos in group theory.
Still, we cannot answer some of the simplest questions.

Richard Brauer

1. DEFINITION OF A GROUP REPRESENTATION

Operations of a group on an arbitrary set were studied in Chapter 5. In this chapter we consider the case that the group elements act as linear operators on a vector space. Such an operation defines a homomorphism from G to the general linear group. A homomorphism to the general linear group is called a *matrix representation.*

The finite rotation groups are good examples to keep in mind. The group T of rotations of a tetrahedron, for example, operates on a three-dimensional space V by rotations. We didn't write down the matrices which represent this action explicitly in Chapter 5; let us do so now. A natural choice of basis has the coordinate axes passing through the midpoints of three of the edges, as illustrated below:



(1.1) **Figure.**

Let $y_i \in T$ denote the rotation by π around an edge, and let $x \in T$ denote rotation by $2\pi/3$ around the front vertex. The matrices representing these operations are

$$(1.2)\qquad R_{y_1} = \begin{bmatrix} 1 & & \\ & -1 & \\ & & -1 \end{bmatrix},\quad R_{y_2} = \begin{bmatrix} -1 & & \\ & 1 & \\ & & -1 \end{bmatrix},\quad R_{y_3} = \begin{bmatrix} -1 & & \\ & -1 & \\ & & 1 \end{bmatrix},$$
$$R_x = \begin{bmatrix} & & 1 \\ 1 & & \\ & 1 & \end{bmatrix}.$$

The rotations $\{y_i, x\}$ generate the group T, and the matrices $\{R_{y_i}, R_x\}$ generate an isomorphic group of matrices.

It is also easy to write down matrices which represent the actions of C_n, D_n, and O explicitly, but I is fairly complicated.

An n-dimensional *matrix representation* of a group G is a homomorphism

$$(1.3)\qquad R: G \longrightarrow GL_n(F),$$

where F is a field. We will use the notation R_g for the image of g. So each R_g is an invertible matrix, and multiplication in G carries over to matrix multiplication; that is, $R_{gh} = R_g R_h$. The matrices (1.2) describe a three-dimensional matrix representation of T. It happens to be *faithful,* meaning that R is an injection and therefore maps T isomorphically to its image, a subgroup of $GL_3(\mathbb{R})$. Matrix representations are not required to be faithful.

When we study representations, it is essential to work as much as possible without fixing a basis, and to facilitate this, we introduce the concept of a representation of a group on a finite-dimensional vector space V. We denote by

$$(1.4)\qquad GL(V)$$

the group of invertible linear operators on V, the multiplication law being, as always, composition of functions. The choice of a basis of V defines an isomorphism of this group with the group of invertible matrices:

$$(1.5)\qquad \begin{aligned} GL(V) &\longrightarrow GL_n(F) \\ T &\rightsquigarrow \text{matrix of } T. \end{aligned}$$

By a *representation of G on V*, we mean a homomorphism

$$(1.6)\qquad \rho: G \longrightarrow GL(V).$$

The *dimension* of the representation ρ is defined to be the dimension of the vector space V. We will study only representations on *finite-dimensional* vector spaces.

Matrix representations can be thought of as representations of G on the space F^n of column vectors.

Let ρ be a representation. We will denote the image of an element g in $GL(V)$ by ρ_g. Thus ρ_g is a linear operator on V, and $\rho_{gh} = \rho_g \rho_h$. If a basis $\mathbf{B} = (v_1, \dots, v_n)$

is given, the representation ρ defines a matrix representation R by the rule

$$R_g = \text{matrix of } \rho_g. \tag{1.7}$$

We may write this matrix symbolically, as in Chapter 4 (3.1), as

$$\rho_g(\mathbf{B}) = \mathbf{B}R_g. \tag{1.8}$$

If X is the coordinate vector of a vector $v \in V$, that is, if $v = \mathbf{B}X$, then

$$R_g X \text{ is the coordinate vector of } \rho_g(v). \tag{1.9}$$

The rotation groups are examples of representations on a real vector space V without regard to a choice of basis. The rotations are linear operators in $GL(V)$. In (1.1) we chose a basis for V, thereby realizing the elements of T as the matrices (1.2) and obtaining a matrix representation.

So all representations of G on finite-dimensional vector spaces can be reduced to matrix representations if we are willing to choose a basis. We may need to choose one in order to make explicit calculations, but then we must study what happens when we change our basis, which properties are independent of the choice of basis, and which choices are the good ones.

A change of basis in V given by a matrix P changes a matrix representation R to a *conjugate representation* $R' = PRP^{-1}$, that is,

$$R_g' = PR_g P^{-1} \quad \text{for every } g. \tag{1.10}$$

This follows from rule (3.4) in Chapter 4 for change of basis.

There is an equivalent concept, namely that of *operation* of a group G on a vector space V. When we speak of an operation on a vector space, we always mean one which is compatible with the vector space structure—otherwise we shouldn't be thinking of V as a vector space. So such an operation is a group operation in the usual sense [Chapter 5 (5.1)]:

$$1v = v \quad \text{and} \quad (gh)v = g(hv), \tag{1.11}$$

for all $g, h \in G$ and all $v \in V$. In addition, every group element is required to act on V as a *linear operator*. Writing out what this means, we obtain the rules

$$g(v + v') = gv + gv' \quad \text{and} \quad g(cv) = cgv \tag{1.12}$$

which, when added to (1.11), give a complete list of axioms for an operation of G on the vector space V. Since G does operate on the underlying set of V, we can speak of orbits and stabilizers as before.

The two concepts "operation of G on V" and "representation of G on V" are equivalent for the same reason that an operation of a group G on a set S is equivalent to a permutation representation (Chapter 5, Section 8): Given a representation ρ of G on V, we define an operation by the rule

$$gv = \rho_g(v), \tag{1.13}$$

and conversely, given an operation, the same formula can be used to define the operator ρ_g for all $g \in G$. It is a linear operator because of (1.12), and the associative law (1.11) shows that $\rho_g \rho_h = \rho_{gh}$.

Thus we have two notations (1.13) for the action of g on v, and we will use them interchangeably. The notation gv is more compact, so we use it when possible.

In order to focus our attention, and because they are the easiest to handle, we will concentrate on *complex* representations for the rest of this chapter. Therefore the vector spaces V which occur are to be interpreted as complex vector spaces, and GL_n will denote the complex general linear group $GL_n(\mathbb{C})$. Every real matrix representation, such as the three-dimensional representation (1.2) of the rotation group T, can be used to define a complex representation, simply by interpreting the real matrices as complex matrices. We will do this without further comment.

2. G-INVARIANT FORMS AND UNITARY REPRESENTATIONS

A matrix representation $R: G \longrightarrow GL_n$ is called *unitary* if all the matrices R_g are unitary, that is, if the image of the homomorphism R is contained in the unitary group. In other words, a unitary representation is a homomorphism

(2.1)
$$R: G \longrightarrow U_n$$

from G to the unitary group.

In this section we prove the following remarkable fact about representations of finite groups.

(2.2) **Theorem.**

(a) Every finite subgroup of GL_n is conjugate to a subgroup of U_n.

(b) Every matrix representation $R: G \longrightarrow GL_n$ of a finite group G is conjugate to a unitary representation. In other words, given R, there is a matrix $P \in GL_n$ such that $PR_gP^{-1} \in U_n$ for every $g \in G$.

(2.3) **Corollary.**

(a) Let A be an invertible matrix of finite order in GL_n, that is, such that $A^r = I$ for some r. Then A is diagonalizable: There is a $P \in GL_n$ so that PAP^{-1} is diagonal.

(b) Let $R: G \longrightarrow GL_n$ be a representation of a finite group G. Then for every $g \in G$, R_g is a diagonalizable matrix.

Proof of the corollary. (a) The matrix A generates a finite subgroup of GL_n. By Theorem (2.2), this subgroup is conjugate to a subgroup of the unitary group. Hence A is conjugate to a unitary matrix. The Spectral Theorem for normal operators [Chapter 7 (7.3)] tells us that every unitary matrix is diagonalizable. Hence A is diagonalizable.

(b) The second part of the corollary follows from the first, because every element g of a finite group has finite order. Since R is a homomorphism, R_g has finite order too. □

The two parts of Theorem (2.2) are more or less the same. We can derive (a) from (b) by considering the inclusion map of a finite subgroup into GL_n as a matrix representation of the group. Conversely, (b) follows by applying (a) to the image of R.

In order to prove part (b), we restate it in basis-free terminology. Consider a hermitian vector space V (a complex vector space together with a positive definite hermitian form $\langle \, , \rangle$). A linear operator T on V is unitary if $\langle v, w \rangle = \langle T(v), T(w) \rangle$ for all $v, w \in V$ [Chapter 7 (5.2)]. Therefore it is natural to call a representation $\rho: G \longrightarrow GL(V)$ *unitary* if ρ_g is a unitary operator for all $g \in G$, that is, if

$$\langle v, w \rangle = \langle \rho_g(v), \rho_g(w) \rangle, \tag{2.4}$$

for all $v, w \in V$ and all $g \in G$. The matrix representation R (1.7) associated to a unitary representation ρ will be unitary in the sense of (2.1), provided that the basis is orthonormal. This follows from Chapter 7 (5.2b).

To simplify notation, we will write condition (2.4) as

$$\langle v, w \rangle = \langle gv, gw \rangle. \tag{2.5}$$

We now turn this formula around and view it as a condition on the form instead of on the operation. Given a representation ρ of G on a vector space V, a form $\langle \, , \rangle$ on V is called *G-invariant* if (2.4), or equivalently, (2.5) holds.

(2.6) **Theorem.** Let ρ be a representation of a finite group G on a complex vector space V. There exists a G-invariant, positive definite hermitian form $\langle \, , \rangle$ on V.

Proof. We start with an arbitrary positive definite hermitian form on V; say we denote it by $\{ \, , \}$. We will use this form to define a G-invariant form, by *averaging over the group*. Averaging over G is a general method which will be used again. It was already used in Chapter 5 (3.2) to find a fixed point of a finite group operation on the plane. The form $\langle \, , \rangle$ we want is defined by the rule

$$\langle v, w \rangle = \frac{1}{N} \sum_{g \in G} \{gv, gw\}, \tag{2.7}$$

where $N = |G|$ is the order of G. The normalization factor $1/N$ is customary but unimportant. Theorem (2.6) follows from this lemma:

(2.8) **Lemma.** The form (2.7) is hermitian, positive definite, and G-invariant.

Proof. The verification of the first two properties is completely routine. For example,

$$\{gv, g(w + w')\} = \{gv, gw + gw'\} = \{gv, gw\} + \{gv, gw'\}.$$

Therefore

$$\langle v, w + w' \rangle = \frac{1}{N} \sum_{g \in G} \{gv, g(w + w')\} = \frac{1}{N} \sum_{g \in G} \{gv, gw\} + \frac{1}{N} \sum_{g \in G} \{gv, gw'\}$$
$$= \langle v, w \rangle + \langle v, w' \rangle.$$

To show that the form $\langle \, , \rangle$ is G-invariant, let g_0 be an element of G. We must show that $\langle g_0 v, g_0 w \rangle = \langle v, w \rangle$ for all $v, w \in V$. By definition,

$$\langle g_0 v, g_0 w \rangle = \frac{1}{N} \sum_{g \in G} \{g g_0 v, g g_0 w\}.$$

There is an important trick for analyzing such a summation, based on the fact that right multiplication by g_0 is a bijective map $G \longrightarrow G$. As g runs over the group, the products gg_0 do too, in a different order. We change notation, substituting g' for gg_0. Then in the sum, g' runs over the group. So we may as well write the sum as being over $g' \in G$ rather than over $g \in G$. This merely changes the order in which the sum is taken. Then

$$\langle g_0 v, g_0 w \rangle = \frac{1}{N} \sum_{g \in G} \{g g_0 v, g g_0 w\} = \frac{1}{N} \sum_{g' \in G} \{g' v, g' w\} = \langle v, w \rangle,$$

as required. Please think this reindexing trick through and understand it. □

Theorem (2.2) follows easily from Theorem (2.6). Any homomorphism $R: G \longrightarrow GL_n$ is the matrix representation associated to a representation (with $V = \mathbb{C}^n$ and $\mathbf{B} = \mathbf{E}$). By Theorem (2.6), there is a G-invariant form $\langle \, , \rangle$ on V, and we choose an orthonormal basis for V with respect to this form. The matrix representation R' obtained via this basis is conjugate to R (1.10) and unitary [Chapter 7 (5.2)]. □

(2.9) **Example.** The matrix $A = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}$ has order 3, and therefore it defines a matrix representation $\{I, A, A^2\}$ of the cyclic group G of order 3. The averaging process (2.7) will produce a G-invariant form from the standard hermitian product X^*Y on $\mathbb{C}^2$. It is

(2.10) $$\langle X, Y \rangle = \frac{1}{3}[X^*Y + (AX)^*(AY) + (A^2X)^*(A^2Y)] = X^*BX,$$

where

(2.11) $$B = \frac{1}{3}[I + A^*A + (A^2)^*(A^2)] = \frac{2}{3}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

3. COMPACT GROUPS

A linear group is called *compact* if it is a closed and bounded subset of the space of matrices [Appendix (3.8)]. The most important compact groups are the orthogonal and unitary groups:

(3.1) **Proposition.** The orthogonal and unitary groups are compact.

Proof. The columns of an orthogonal matrix P form an orthonormal basis, so they have length 1. Hence all of the matrix entries have absolute value ≤ 1. This shows that O_n is contained in the box defined by the inequalities $|p_{ij}| \leq 1$. So it is a bounded set. Because it is defined as the common zeros of a set of continuous functions, it is closed too, hence compact. The proof for the unitary group is the same. □

The main theorems (2.2, 2.6) of Section 2 carry over to compact linear groups without major change. We will work out the case of the circle group $G = SO_2$ as an example. The rotation of the plane through the angle θ was denoted by ρ_θ in Chapter 5. Here we will consider an arbitrary representation of G. To avoid confusion, we denote the element

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \in SO_2 \tag{3.2}$$

by its angle θ, rather than by ρ_θ. Formula (3.2) defines a particular matrix representation of our group, but there are others.

Suppose we are given a *continuous* representation σ of G on a finite-dimensional space V, not necessarily the representation (3.2). Since the group law is addition of angles, the rule for working with σ is $\sigma_{\theta+\eta} = \sigma_\theta\sigma_\eta$. To say that the operation is continuous means that if we choose a basis for V, thereby representing the operation of θ on V by some matrix S_θ, then the entries of S are continuous functions of θ.

Let us try to copy the proof of (2.6). To average over the infinite group G, we replace summation by an integral. We choose any positive definite hermitian form $\{,\}$ on V and define a new form by the rule

$$\langle v, w\rangle = \frac{1}{2\pi}\int_0^{2\pi} \{\sigma_\theta v, \sigma_\theta w\}\, d\theta. \tag{3.3}$$

This form has the required properties. To check G-invariance, fix any element $\theta_0 \in G$, and let $\eta = \theta + \theta_0$. Then $d\eta = d\theta$. Hence

$$\begin{aligned} \langle \sigma_{\theta_0} v, \sigma_{\theta_0} w\rangle &= \frac{1}{2\pi}\int_0^{2\pi} \{\sigma_\theta\sigma_{\theta_0} v, \sigma_\theta\sigma_{\theta_0} w\}\, d\theta \\ &= \frac{1}{2\pi}\int_0^{2\pi} \{\sigma_\eta v, \sigma_\eta w\}\, d\eta = \langle v, w\rangle, \end{aligned} \tag{3.4}$$

as required.

We will not carry the proof through for general groups because some serious work has to be done to find a suitable volume element analogous to $d\theta$ in a given compact group G. In the computation (3.4), it is crucial that $d\theta = d(\theta + \theta_0)$, and we were lucky that the obvious integral was the one to use.

For any compact group G there is a volume element dg called *Haar measure*, which has the property of being translation invariant: If $g_0 \in G$ is fixed and

$g' = gg_0$, then

$$dg = dg'. \tag{3.5}$$

Using this measure, the proof carries over. We will not prove the existence of a Haar measure, but assuming one exists, the same reasoning as in (2.8) proves the following analogue of (2.6) and (2.2):

(3.6) **Corollary.** Let G be a compact subgroup of GL_n. Then

(a) Let σ be a representation of G on a finite-dimensional vector space V. There is a G-invariant, positive definite hermitian form $\langle\, ,\rangle$ on V.

(b) Every continuous matrix representation R of G is conjugate to a unitary representation.

(c) Every compact subgroup G of GL_n is conjugate to a subgroup of U_n. □

4. G-INVARIANT SUBSPACES AND IRREDUCIBLE REPRESENTATIONS

Given a representation of a finite group G on a vector space V, Corollary (2.3) tells us that for each group element g there is a basis of V so that the matrix of the operator ρ_g is diagonal. Obviously, it would be very convenient to have a single basis which would diagonalize ρ_g for all group elements g at the same time. But such a basis doesn't exist very often, because any two diagonal matrices commute with each other. In order to diagonalize the matrices of all ρ_g at the same time, these operators must commute. It follows that any group G which has a faithful representation by diagonal matrices is abelian. We will see later (Section 8) that the converse is also true. If G is a finite abelian group, then every matrix representation R of G is diagonalizable; that is, there is a single matrix P so that PR_gP^{-1} is diagonal for all $g \in G$. In this section we discuss what can be done for finite groups in general.

Let ρ be a representation of a group G on a vector space V. A subspace of V is called *G-invariant* if

$$gw \in W, \quad \textit{for all } w \in W \textit{ and } g \in G. \tag{4.1}$$

So the operation by every group element g must carry W to itself, that is, $gW \subset W$. This is an extension of the concept of T-invariant subspace introduced in Section 3 of Chapter 4. In a representation, the elements of G represent linear operators on V, and we ask that W be an invariant subspace for each of these operators. If W is G-invariant, the operation of G on V will restrict to an operation on W.

As an example, consider the three-dimensional representation of the dihedral group defined by the symmetries of an n-gon Δ [Chapter 5 (9.1)]. So $G = D_n$. There are two proper G-invariant subspaces: The plane containing Δ and the line perpendicular to Δ. On the other hand, there is no proper T-invariant subspace for the representation (1.2) of the tetrahedral group T, because there is no line or plane which is carried to itself by *every* element of T.

If a representation ρ of a group G on a nonzero vector space V has no proper G-invariant subspace, it is called an *irreducible* representation. If there is a proper invariant subspace, ρ is said to be *reducible*. The standard three-dimensional representation of T is irreducible.

When V is the direct sum of G-invariant subspaces: $V = W_1 \oplus W_2$, the representation ρ on V is said to be the *direct sum* of its restrictions ρ_i to W_i, and we write

$$\rho = \rho_1 \oplus \rho_2. \tag{4.2}$$

Suppose this is the case. Choose bases $\mathbf{B}_1, \mathbf{B}_2$ of W_1, W_2, and let $\mathbf{B} = (\mathbf{B}_1, \mathbf{B}_2)$ be the basis of V obtained by listing these two bases in order [Chapter 3 (6.6)]. Then the matrix R_g of ρ_g will have the block form

$$R_g = \left[\begin{array}{c|c} A_g & 0 \\ \hline 0 & B_g \end{array}\right], \tag{4.3}$$

where A_g is the matrix of ρ_{1g} with respect to $\mathbf{B}_1$ and B_g is the matrix of ρ_{2g} with respect to $\mathbf{B}_2$. Conversely, if the matrices R_g have such a block form, then the representation is a direct sum.

For example, consider the rotation group $G = D_n$ operating on $\mathbb{R}^3$ by symmetries of an n-gon Δ. If we choose an orthonormal basis $\mathbf{B}$ so that v_1 is perpendicular to the plane of Δ and v_2 passes through a vertex, then the rotations corresponding to our standard generators x, y [Chapter 5 (3.6)] are represented by the matrices

$$R_x = \begin{bmatrix} 1 & & \\ & c_n & -s_n \\ & s_n & c_n \end{bmatrix}, \quad R_y = \begin{bmatrix} -1 & & \\ & 1 & \\ & & -1 \end{bmatrix}, \tag{4.4}$$

where $c_n = \cos(2\pi/n)$ and $s_n = \sin(2\pi/n)$. So R is a direct sum of a one-dimensional representation A,

$$A_x = [1], \quad A_y = [-1], \tag{4.5}$$

and a two-dimensional representation B,

$$B_x = \begin{bmatrix} c_n & -s_n \\ s_n & c_n \end{bmatrix}, \quad B_y = \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}. \tag{4.6}$$

The representation B is the basic two-dimensional representation of D_n as symmetries of Δ in the plane.

On the other hand, even if a representation ρ is reducible, the matrices R_g will not have a block form unless the given basis for V is compatible with the direct sum decomposition. Until we have made a further analysis, it will be difficult to tell that a representation is reducible, when it is presented using the wrong basis.

(4.7) **Proposition.** Let ρ be a unitary representation of G on a hermitian vector space V, and let W be a G-invariant subspace. The orthogonal complement $W^{\perp}$ is also G-invariant, and ρ is a direct sum of its restrictions to W and $W^{\perp}$.

Proof. Let $v \in W^{\perp}$, so that $v \perp W$. Since the operators ρ_g are unitary, they preserve orthogonality [Chapter 7 (5.2)], so $gv \perp gW$. Since W is G-invariant, $W = gW$, so $gv \perp W$. Therefore $gv \in W^{\perp}$. This shows that $W^{\perp}$ is G-invariant. We know that $V = W \oplus W^{\perp}$ by Chapter 7 (2.7). □

This proposition allows us to decompose a representation as a direct sum, provided that there is a proper invariant subspace. Together with induction, this gives us the following corollary:

(4.8) **Corollary.** Every unitary representation $\rho: G \longrightarrow GL(V)$ on a hermitian vector space V is a direct sum of irreducible representations. □

Combining this corollary with (2.2), we obtain the following:

(4.9) **Corollary.** *Maschke's Theorem:* Every representation of a finite group G is a direct sum of irreducible representations. □

5. CHARACTERS

Two representations $\rho: G \longrightarrow GL(V)$ and $\rho': G \longrightarrow GL(V')$ of a group G are called *isomorphic,* or *equivalent,* if there is an isomorphism of vector spaces $T: V \longrightarrow V'$ which is compatible with the operation of G:

$$gT(v) = T(gv) \quad \text{or} \quad \rho_g' T(v) = T(\rho_g(v)), \tag{5.1}$$

for all $v \in V$ and $g \in G$. If $\mathbf{B}$ is a basis for V and if $\mathbf{B}' = T(\mathbf{B})$ is the corresponding basis of V', then the associated matrix representations R_g and R_g' will be *equal.*

For the next four sections, we restrict our attention to representations of finite groups. We will see that there are relatively few isomorphism classes of irreducible representations of a finite group. However, each representation has a complicated description in terms of matrices. The secret to understanding representations is not to write down the matrices explicitly unless absolutely necessary. So to facilitate classification we will throw out most of the information contained in a representation ρ, keeping only an essential part. What we will work with is the trace, called the *character,* of ρ. Characters are usually denoted by χ.

The *character* χ of a representation ρ is the map $\chi: G \longrightarrow \mathbb{C}$ defined by

$$\chi(g) = \text{trace}(\rho_g). \tag{5.2}$$

If R is the matrix representation obtained from ρ by a choice of basis for V, then

$$\chi(g) = \text{trace}(R_g) = \lambda_1 + \cdots + \lambda_n, \tag{5.3}$$

where λ_i are the eigenvalues of R_g, or of ρ_g.

The *dimension* of a character χ is defined to be the dimension of the representation ρ. The character of an irreducible representation is called an *irreducible character.*

Here are some basic properties of the character:

(5.4) **Proposition.** Let χ be the character of a representation ρ of a finite group G on a vector space V.

(a) $\chi(1)$ is the dimension of the character [the dimension of V].
(b) $\chi(g) = \chi(hgh^{-1})$ for all $g, h \in G$. In other words, the character is constant on each conjugacy class.
(c) $\chi(g^{-1}) = \overline{\chi(g)}$ [the complex conjugate of $\chi(g)$].
(d) If χ' is the character of another representation ρ', then the character of the direct sum $\rho \oplus \rho'$ is $\chi + \chi'$.

Proof. The symbol 1 in assertion (a) denotes the identity element of G. This property is trivial: $\chi(1) = \text{trace } I = \dim V$. Property (b) is true because the matrix representation R associated to ρ is a homomorphism, which shows that $R_{hgh^{-1}} = R_h R_g R_h^{-1}$, and because trace $(R_h R_g R_h^{-1}) =$ trace R_g [Chapter 4 (4.18)]. Property (d) is also clear, because the trace of the block matrix (4.3) is the sum of the traces of A_g and B_g.

Property (c) is less obvious. If the eigenvalues of R_g are $\lambda_1, \ldots, \lambda_n$, then the eigenvalues of $R_{g^{-1}} = (R_g)^{-1}$ are $\lambda_1^{-1}, \ldots, \lambda_n^{-1}$. The assertion of (c) is

$$\chi(g^{-1}) = \lambda_1^{-1} + \cdots + \lambda_n^{-1} = \overline{\lambda}_1 + \cdots + \overline{\lambda}_n = \overline{\chi(g)},$$

and to show this we use the fact that G is a finite group. Every element g of G has finite order. If $g^r = 1$, then R_g is a matrix of order r, so its eigenvalues $\lambda_1, \ldots, \lambda_n$ are roots of unity. This implies that $|\lambda_i| = 1$, hence that $\lambda_i^{-1} = \overline{\lambda}_i$ for each i. □

In order to avoid confusing cyclic groups with conjugacy classes, we will denote conjugacy classes by the roman letter C, rather than an italic C, in this chapter. Thus the conjugacy class of an element $g \in G$ will be denoted by C_g.

We shall note two things which simplify the computation of a character. First of all, since the value of χ depends only on the conjugacy class of an element $g \in G$ (5.4b), we need only determine the values of χ on one representative element in each class. Second, since the value of the character $\chi(g)$ is the trace of the operator ρ_g and since the trace doesn't depend on the choice of a basis, we are free to choose a convenient one. Moreover, we may select a convenient basis for each individual group element. There is no need to use the same basis for all elements.

As an example, let us determine the character χ of the rotation representation of the tetrahedral group T defined by (1.2). There are four conjugacy classes in T, and they are represented by the elements $1, x, x^2, y$, where as before x is a rotation by $2\pi/3$ about a vertex and y is a rotation by π about the center of an edge. The values of the character on these representatives can be read off from the matrices (1.2):

$$\chi(1) = 3, \quad \chi(x) = 0, \quad \chi(x^2) = 0, \quad \chi(y) = -1. \tag{5.5}$$

It is sometimes useful to think of a character χ as a vector. We can do this by

listing the elements of G in some order: $G = \{g_1, \ldots, g_N\}$; then the vector representing χ will be

$$\chi = (\chi(g_1), \ldots, \chi(g_N))^t. \tag{5.6}$$

Since χ is constant on conjugacy classes, it is natural to list G by listing the conjugacy classes and then running through each conjugacy class in some order. If we do this for the character (5.5), listing C_1, C_x, C_{x^2}, C_y in that order, the vector we obtain is

$$\chi = (3; 0, 0, 0, 0; 0, 0, 0, 0; -1, -1, -1)^t. \tag{5.7}$$

We will not write out such a vector explicitly again.

The main theorem on characters relates them to the hermitian dot product on $\mathbb{C}^N$. This is one of the most beautiful theorems of algebra, both because its statement is intrinsically so elegant and because it simplifies the problem of classifying representations so much. We define

$$\langle \chi, \chi' \rangle = \frac{1}{N} \sum_g \overline{\chi(g)} \chi'(g), \tag{5.8}$$

where $N = |G|$. If χ, χ' are represented by vectors as in (5.7), this is the standard hermitian product, renormalized by the factor $1/N$.

(5.9) **Theorem.** Let G be a group of order N, let $\rho_1, \rho_2, \ldots$ represent the distinct isomorphism classes of irreducible representations of G, and let χ_i be the character of ρ_i.

(a) *Orthogonality Relations:* The characters χ_i are orthonormal. In other words, $\langle \chi_i, \chi_j \rangle = 0$ if $i \neq j$, and $\langle \chi_i, \chi_i \rangle = 1$ for each i.

(b) There are finitely many isomorphism classes of irreducible representations, the same number as the number of conjugacy classes in the group.

(c) Let d_i be the dimension of the irreducible representation ρ_i, and let r be the number of irreducible representations. Then d_i divides N, and

$$N = d_1^2 + \cdots + d_r^2. \tag{5.10}$$

This theorem will be proved in Section 9, with the exception of the assertion that d_i divides N, which we will not prove.

A complex-valued function $\varphi: G \longrightarrow \mathbb{C}$ which is constant on each conjugacy class is called a *class function*. Since a class function is constant on each class, it may also be described as a function on the set of conjugacy classes. The class functions form a complex vector space, which we denote by $\mathscr{C}$. We use the form defined by (5.8) to make $\mathscr{C}$ into a hermitian space.

(5.11) **Corollary.** The irreducible characters form on orthonormal basis of $\mathscr{C}$.

This follows from (5.9a,b). The characters are linearly independent because they are orthogonal, and they span because the dimension of $\mathscr{C}$ is the number of conjugacy classes, which is r. □

The corollary allows us to decompose a given character as a linear combination of the irreducible characters, using the formula for orthogonal projection [Chapter 7 (3.8)]. For let χ be the character of a representation ρ. By Corollary (4.9), ρ is isomorphic to a direct sum of the irreducible representations $\rho_1,\ldots,\rho_r$; say we write this symbolically as $\rho = n_1\rho_1 \oplus \cdots \oplus n_r\rho_r$, where n_i are nonnegative integers and where $n\rho$ stands for the direct sum of n copies of the representation ρ. Then $\chi = n_1\chi_1 + \cdots + n_r\chi_r$. Since $(\chi_1,\ldots,\chi_r)$ is an orthonormal basis, we have the following:

(5.12) **Corollary.** Let $\chi_1,\ldots,\chi_r$ be the irreducible characters of a finite group G, and let χ be any character. Then $\chi = n_1\chi_1 + \cdots + n_r\chi_r$, where $n_i = \langle \chi, \chi_i \rangle$. □

(5.13) **Corollary.** If two representations ρ, ρ' have the same character, they are isomorphic.

For let χ, χ' be the characters of two representations ρ, ρ', where $\rho = n_1\rho_1 \oplus \cdots \oplus n_r\rho_r$ and $\rho' = n_1'\rho_1 \oplus \cdots \oplus n_r'\rho_r$. Then the characters of these representations are $\chi = n_1\chi_1 + \cdots + n_r\chi_r$ and $\chi' = n_1'\chi_1 + \cdots + n_r'\chi_r$. Since $\chi_1,\ldots,\chi_r$ are linearly independent, $\chi = \chi'$ implies that $n_i = n_i'$ for each i. □

(5.14) **Corollary.** A character χ has the property $\langle \chi, \chi \rangle = 1$ if and only if it is irreducible.

For if $\chi = n_1\chi_1 + \cdots + n_r\chi_r$, then $\langle \chi, \chi \rangle = n_1^2 + \cdots + n_r^2$. This gives the value 1 if and only if a single n_i is 1 and the rest are zero. □

The evaluation of $\langle \chi, \chi \rangle$ is a very practical way to check irreducibility of a representation. For example, let χ be the character (5.7) of the representation (1.2). Then $\langle \chi, \chi \rangle = (3^2+1+1+1)/12 = 1$. So χ is irreducible.

Part (c) of Theorem (5.9) should be contrasted with the Class Equation [Chapter 6 (1.7)]. Let $C_1,\ldots,C_r$ be the conjugacy classes in G, and let $c_i = |C_i|$ be the order of the conjugacy class. Then c_i divides N, and $N = c_1 + \cdots + c_r$. Though there is the same number of conjugacy classes as irreducible representations, their exact relationship is very subtle.

As our first example, we will determine the irreducible representations of the dihedral group D_3 [Chapter 5 (3.6)]. There are three conjugacy classes, $C_1 = \{1\}$, $C_2 = \{y, xy, x^2y\}$, $C_3 = \{x, x^2\}$ [Chapter 6 (1.8)], and therefore three irreducible representations. The only solution of equation (5.10) is $6 = 1^2 + 1^2 + 2^2$, so D_3 has two one-dimensional representations ρ_1,ρ_2 and one irreducible two-dimensional representation ρ_3. Every group G has the *trivial* one-dimensional representation

($R_g = 1$ for all g); let us call it ρ_1. The other one-dimensional representation is the *sign representation* of the symmetric group S_3, which is isomorphic to D_3: $R_g = \text{sign}(g) = \pm 1$. This is the representation (4.5); let us call it ρ_2. The two-dimensional representation is defined by (4.6); call it ρ_3.

Rather than listing the characters χ_i as vectors, we usually assemble them into a *character table.* In this table, the three conjugacy classes are represented by the elements $1, y, x$. The orders of the conjugacy classes are given above them. Thus $|C_y| = 3$.

		conjugacy class			
		(1)	(3)	(2)	order of the class
		1	y	x	representative element
irreducible character	χ_1	1	1	1	value of the character
	χ_2	1	-1	1	
	χ_3	2	0	-1	

(5.15) CHARACTER TABLE FOR D_3

In such a table, the top row, corresponding to the trivial character, consists entirely of 1's. The first column contains the dimensions of the representations, because $\chi_i(1) = \dim \rho_i$.

To evaluate the bilinear form (5.8) on the characters, remember that there are three elements in the class of y and two in the class of x. Thus

$$\langle \chi_3, \chi_3 \rangle = \frac{1}{N} \sum_g \overline{\chi_3(g)}\chi_3(g) = (1 \cdot (\overline{\chi_3(1)}\chi_3(1)) + 3 \cdot (\overline{\chi_3(y)}\chi_3(y)) + 2 \cdot (\overline{\chi_3(x)}\chi_3(x)))/6$$

$$= (1 \cdot \overline{2} \cdot 2 + 3 \cdot \overline{0} \cdot 0 + 2 \cdot (\overline{-1}) \cdot (-1))/6 = 1.$$

This confirms the fact that ρ_3 is irreducible. □

As another example, consider the cyclic group $C_3 = \{1, x, x^2\}$ of order 3. Since C_3 is abelian, there are three conjugacy classes, each consisting of one element. Theorem (5.9) shows that there are three irreducible representations, and that each has dimension 1. Let $\zeta = \frac{1}{2}(-1 + \sqrt{3}i)$ be a cube root of 1. The three representations are

(5.16) $$\rho_{1_x} = 1, \quad \rho_{2_x} = \zeta, \quad \rho_{3_x} = \zeta^2.$$

	1	x	x^2
χ_1	1	1	1
χ_2	1	ζ	ζ^2
χ_3	1	ζ^2	ζ

(5.17) CHARACTER TABLE FOR C_3

Note that $\overline{\zeta} = \zeta^2$. So

$$\langle \chi_2, \chi_3 \rangle = (\bar{1} \cdot 1 + \bar{\zeta}\zeta^2 + \bar{\zeta}^2\zeta)/3 = (1 + \zeta + \zeta^2)/3 = 0,$$

which agrees with the orthogonality relations.

As a third example, let us determine the character table of the tetrahedral group T. The conjugacy classes C_1, C_x, C_{x^2}, C_y were determined above, and the Class Equation is $12 = 1+4+4+3$. The only solution of (5.10) is $12 = 1^2+1^2+1^2+3^2$, so there are four irreducible representations, of dimensions 1, 1, 1, 3. Now it happens that T has a normal subgroup H of order 4 which is isomorphic to the Klein four group, and such that the quotient $\overline{T} = T/H$ is cyclic of order 3. Any representation $\overline{\rho}$ of $\overline{T}$ will give a representation of T by composition:

$$T \xrightarrow{\pi} \overline{T} \xrightarrow{\overline{\rho}} GL(V).$$

Thus the three one-dimensional representations of the cyclic group determine representations of T. Their characters χ_1, χ_2, χ_3 can be determined from (5.17). The character (5.5) is denoted by χ_4 in the table below.

	(1) 1	(4) x	(4) x^2	(3) y
χ_1	1	1	1	1
χ_2	1	ζ	ζ^2	1
χ_3	1	ζ^2	ζ	1
χ_4	3	0	0	-1

(5.18) CHARACTER TABLE FOR T

Various properties of the group can be read off easily from the character table. Let us forget that this is the character table for T, and suppose that it has been given to us as the character table of an unknown group G. After all, it is conceivable that another isomorphism class of groups has the same characters.

The order of G is 12, the sum of the orders of the conjugacy classes. Next, since the dimension of ρ_2 is 1, $\chi_2(y)$ is the trace of the 1×1 matrix ρ_{2_y}. So the fact that $\chi_2(y) = 1$ shows that $\rho_{2_y} = 1$ too, that is, that y is in the kernel of ρ_2. In fact, the kernel of ρ_2 is identified as the union of the two conjugacy classes $C_1 \cup C_y$. This is a subgroup H of order 4 in G. Moreover, H is the Klein four group. For if H were C_4, its unique element of order 2 would have to be in a conjugacy class by itself. It also follows from the value of $\chi_2(x)$ that the order of x is divisible by 3. Going back to our list [Chapter 6 (5.1)] of groups of order 12, we see that $G \approx T$.

6. PERMUTATION REPRESENTATIONS AND THE REGULAR REPRESENTATION

Let S be a set. We can construct a representation of a group G from an operation of G on S, by passing to the vector space $V = V(S)$ of formal linear combinations [Chapter 3 (3.21)]

$$v = \sum_i a_i s_i, \quad a_i \in \mathbb{C}.$$

An element $g \in G$ operates on vectors by permuting the elements of S, leaving the coefficients alone:

(6.1)
$$gv = \sum_i a_i g s_i .$$

If we choose an ordering $s_1, \ldots, s_n$ of S and take the basis $(s_1, \ldots, s_n)$ for V, then R_g is the permutation matrix which describes the operation of g on S.

For example, let $G = T$ and let S be the set of faces of the tetrahedron: $S = (f_1, \ldots, f_4)$. The operation of G on S defines a four-dimensional representation of G. Let x denote the rotation by $2\pi/3$ about a face f_1 and y the rotation by π about an edge as before. Then if the faces are numbered appropriately, we will have

(6.2)
$$R_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad R_y = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

We will call ρ (or R) the representation *associated* to the operation of G on S and will often refer to ρ as a *permutation representation*, though that expression has a meaning in another context as well (Chapter 5, Section 8).

If we decompose a set on which G operates into orbits, we will obtain a decomposition of the associated representation as a direct sum. This is clear. But there is an important new feature: The fact that linear combinations are available in $V(S)$ allows us to decompose the representation further. Even though S may consist of a single orbit, the associated permutation representation ρ will *never* be irreducible, unless S has only one element. This is because the vector $w = s_1 + \cdots + s_r$ is fixed by every permutation of the basis, and so the one-dimensional subspace $W = \{cw\}$ is G-invariant. The trivial representation is a summand of every permutation representation.

It is easy to compute the character of a permutation representation:

(6.3)
$$\chi(g) = \text{number of elements of } S \text{ fixed by } g,$$

because for every index fixed by a permutation, there is a 1 on the diagonal of the associated permutation matrix, and the other diagonal entries are 0. For example, the character χ of the representation of T on the faces of a tetrahedron is

(6.4)

	1	x	x^2	y
χ	4	1	1	0

and the character table (5.18) shows that $\chi = \chi_1 + \chi_4$. Therefore $\rho \approx \rho_1 \oplus \rho_4$ by Corollary (5.13). As another example, the character of the operation of T on the six edges of the tetrahedron is

(6.5)

	1	x	x^2	y
χ	6	0	0	2

and using (5.18) again, we find that $\chi = \chi_1 + \chi_2 + \chi_3 + \chi_4$.

The *regular representation* ρ^{reg} of G is the representation associated to the op-

eration of G on itself by left multiplication. In other words, we let $S = G$, with the operation of left multiplication. This is not an especially interesting operation, but its associated representation is very interesting. Its character χ^{reg} is particularly simple:

$$\chi^{\text{reg}}(1) = N, \quad \text{and} \quad \chi^{\text{reg}}(g) = 0, \quad \text{if } g \neq 1, \tag{6.6}$$

where $N = G$. The first formula is clear: $\chi(1) = \dim \rho$ for any representation ρ, and ρ^{reg} has dimension N. The second follows from (6.3), because multiplication by g does not fix any element of G, unless $g = 1$.

Because of this formula, it is easy to compute $\langle \chi^{\text{reg}}, \chi \rangle$ for the character χ of any representation ρ by the orthogonal projection formula (5.12). The answer is

$$\langle \chi^{\text{reg}}, \chi \rangle = \dim \rho, \tag{6.7}$$

because $\chi(1) = \dim \rho$. This allows us to write χ^{reg} as a linear combination of the irreducible characters:

(6.8) **Corollary.** $\chi^{\text{reg}} = d_1\chi_1 + \cdots + d_r\chi_r$, and $\rho^{\text{reg}} \approx d_1\rho_1 \oplus \cdots \oplus d_r\rho_r$, where d_i is the dimension of ρ_i and $d_i\rho_i$ stands for the direct sum of d_i copies of ρ_i. □

Isn't this a nice formula? We can deduce formula (5.10) from (6.8) by counting dimensions. This shows that formula (5.10) of Theorem (5.9) follows from the orthogonality relations.

For instance, for the group D_3, the character of the regular representation is

	1	x	y
χ^{reg}	6	0	0

and Table (5.15) shows that $\chi^{\text{reg}} = \chi_1 + \chi_2 + 2\chi_3$, as expected.

As another example, consider the regular representation R of the cyclic group $\{1, x, x^2\}$ of order 3. The permutation matrix representing x is

$$R_x = \begin{bmatrix} & & 1 \\ 1 & & \\ & 1 & \end{bmatrix}.$$

Its eigenvalues are $1, \zeta, \zeta^2$, where $\zeta = \frac{1}{2}(-1 + \sqrt{3}i)$. Thus R_x is conjugate to

$$R'_x = \begin{bmatrix} 1 & & \\ & \zeta & \\ & & \zeta^2 \end{bmatrix}.$$

This matrix displays the decomposition $\rho^{\text{reg}} \approx \rho_1 \oplus \rho_2 \oplus \rho_3$ of the regular representation into irreducible one-dimensional representations.

7. THE REPRESENTATIONS OF THE ICOSAHEDRAL GROUP

In this section we determine the irreducible characters of the icosahedral group. So far, we have seen only its trivial representation ρ_1 and the representation of dimension 3 as a rotation group. Let us denote the rotation representation by ρ_2. There are

five conjugacy classes in I [Chapter 6 (2.2)], namely

(7.1)
$$\begin{aligned} C_1 &= \{1\}, \\ C_2 &= 15 \text{ rotations "}x\text{" through the angle } \pi, \\ C_3 &= 20 \text{ rotations "}y\text{" by } 2\pi/3,\ 4\pi/3, \\ C_4 &= 12 \text{ rotations "}z\text{" by } 2\pi/5,\ 8\pi/5, \\ C_5 &= 12 \text{ rotations "}z^2\text{" by } 4\pi/5,\ 6\pi/5, \end{aligned}$$

and therefore there are three more irreducible representations. Given what we know already, the only solution to (5.10) is $d_i = 1, 3, 3, 4, 5$:

$$60 = 1^2+3^2+3^2+4^2+5^2.$$

We denote the remaining representations by ρ_3, ρ_4, ρ_5, where $\dim \rho_3 = 3$, and so on. A good way to find the missing irreducible representations is to decompose some known permutation representations. We know that I operates on a set of five elements [Chapter 6 (2.6)]. This gives us a five-dimensional representation ρ'. As we saw in Section 6, the trivial representation is a summand of ρ'. Its orthogonal complement turns out to be the required irreducible four-dimensional representation: $\rho' = \rho_1 \oplus \rho_4$. Also, I permutes the set of six axes through the centers of opposite faces of the dodecahedron. Let the corresponding six-dimensional representation be ρ''. Then $\rho'' = \rho_1 \oplus \rho_5$. We can check this by computing the characters of ρ_4 and ρ_5 and applying Theorem (5.9). The characters χ_4, χ_5 are computed from χ', χ'' by subtracting $\chi_1 = 1$ from each value (5.4d). For example, ρ' realizes x as an even permutation of $\{\mathbf{1}, \ldots, \mathbf{5}\}$ of order 2, so it is a product of two disjoint transpositions, which fixes one index. Therefore $\chi'(x) = 1$, and $\chi_4(x) = 0$.

The second three-dimensional representation ρ_3 is fairly subtle because it is so similar to ρ_2. It can be obtained this way: Since I is isomorphic to A_5, we may view it as a normal subgroup of the symmetric group S_5. Conjugation by an element p of S_5 which is not in A_5 defines an automorphism σ of A_5. This automorphism interchanges the two conjugacy classes C_4, C_5. The other conjugacy classes are not interchanged, because their elements have different orders. For example, in cycle notation, let $z = (\mathbf{12345})$ and let $p = (\mathbf{2354})$. Then $p^{-1}zp = (\mathbf{4532})(\mathbf{12345})(\mathbf{2354}) = (\mathbf{13524}) = z^2$. The representation ρ_3 is $\rho_2 \circ \sigma$.

The character of ρ_3 is computed from that of ρ_2 by interchanging the values for z, z^2. Once these characters are computed, verification of the relations $\langle \chi_i, \chi_j \rangle = 0$, $\langle \chi_i, \chi_i \rangle = 1$ shows that the representations are irreducible and that our list is correct.

	(1) 1	(15) x	(20) y	(12) z	(12) z^2
χ_1	1	1	1	1	1
χ_2	3	-1	0	α	β
χ_3	3	-1	0	β	α
χ_4	4	0	1	-1	-1
χ_5	5	1	-1	0	0

(7.2) CHARACTER TABLE FOR $I = A_5$

In this table, α is the trace of a three-dimensional rotation through the angle $2\pi/5$, which is

$$\alpha = 1 + 2\cos 2\pi/5 = \tfrac{1}{2}(-1 + \sqrt{5}),$$

and β is computed similarly: $\beta = 1 + 2\cos 4\pi/5 = \frac{1}{2}(-1 - \sqrt{5})$.

8. ONE-DIMENSIONAL REPRESENTATIONS

Let ρ be a one-dimensional representation of a group G. So R_g is a 1×1 matrix, and $\chi(g) = R_g$, provided that we identify a 1×1 matrix with its single entry. Therefore in this case the character χ is a homomorphism $\chi\colon G \longrightarrow \mathbb{C}^\times$, that is, it satisfies the rule

$$(8.1) \qquad \chi(gh) = \chi(g)\chi(h), \quad \text{if } \dim \rho = 1.$$

Such a character is called *abelian*. Please note that formula (8.1) is not true for characters of dimension >1.

If G is a finite group, the values taken on by an abelian character χ are always roots of 1:

$$(8.2) \qquad \chi(g)^r = 1$$

for some r, because the element g has finite order.

The one-dimensional characters form a group under multiplication of functions:

$$(8.3) \qquad \chi\chi'(g) = \chi(g)\chi'(g).$$

This group is called the *character group* of G and is often denoted by $\hat{G}$. The character group is especially important when G is abelian, because of the following fact:

(8.4) **Theorem.** If G is a finite abelian group, then every irreducible representation of G is one-dimensional.

Proof. Since G is abelian, every conjugacy class consists of one element. So the number of conjugacy classes is N. By Theorem (5.9), there are N irreducible representations, and $d_1 = d_2 = \cdots = d_r = 1$. □

9. SCHUR'S LEMMA, AND PROOF OF THE ORTHOGONALITY RELATIONS

Let ρ, ρ' be representations of a group G on two vector spaces V, V'. We will call a linear transformation $T\colon V \longrightarrow V'$ *G-invariant* if it is compatible with the two operations of G on V and V', that is, if

$$(9.1) \qquad gT(v) = T(gv), \quad \text{or} \quad \rho_g'(T(v)) = T(\rho_g(v)),$$

for all $g \in G$ and $v \in V$. Thus an isomorphism of representations (Section 5) is a

bijective G-invariant transformation. We could also write (9.1) as

(9.2) $$\rho_g' \circ T = T \circ \rho_g, \quad \text{for all } g \in G.$$

Let bases $\mathbf{B}, \mathbf{B}'$ for V and V' be given, and let R_g, R_g' and A denote the matrices of ρ_g, ρ_g' and T with respect to these bases. Then (9.2) reads

(9.3) $$R_g'A = AR_g, \quad \text{for all } g \in G.$$

The special case that $\rho = \rho'$ is very important. A *G-invariant* linear operator T on V is one which commutes with ρ_g for every $g \in G$:

(9.4) $$\rho_g \circ T = T \circ \rho_g \quad \text{or} \quad R_gA = AR_g.$$

These formulas just repeat (9.2) and (9.3) when $\rho = \rho'$.

(9.5) **Proposition.** *The kernel and image of a G-invariant linear transformation $T: V \to V'$ are G-invariant subspaces of V and V' respectively.*

Proof. The kernel and image of any linear transformation are subspaces. Let us show that $\ker T$ is G-invariant: We want to show that $gv \in \ker T$ if $v \in \ker T$, or that $T(gv) = 0$ if $T(v) = 0$. Well,

$$T(gv) = gT(v) = g0 = 0.$$

Similarly, if $v' \in \operatorname{im} T$, then $v' = T(v)$ for some $v \in V$. Then

$$gv' = gT(v) = T(gv),$$

so $gv' \in \operatorname{im} T$ too. □

(9.6) **Theorem.** *Schur's Lemma:* Let ρ, ρ' be two irreducible representations of G on vector spaces V, V', and let $T: V \longrightarrow V'$ be a G-invariant transformation.

(a) Either T is an isomorphism, or else $T = 0$.
(b) If $V = V'$ and $\rho = \rho'$, then T is multiplication by a scalar.

Proof. (a) Since ρ is irreducible and since $\ker T$ is a G-invariant subspace, $\ker T = V$ or else $\ker T = 0$. In the first case, $T = 0$. In the second case, T is injective and maps V isomorphically to its image. Then $\operatorname{im} T$ is not zero. Since ρ' is irreducible and $\operatorname{im} T$ is G-invariant, $\operatorname{im} T = V'$. Therefore T is an isomorphism.

(b) Suppose $V = V'$, so that T is a linear operator on V. Choose an eigenvalue λ of T. Then $(T - \lambda I) = T_1$ is also G-invariant. Its kernel is nonzero because it contains an eigenvector. Since ρ is irreducible, $\ker T_1 = V$, which implies that $T_1 = 0$. Therefore $T = \lambda I$. □

The averaging process can be used to create a G-invariant transformation from any linear transformation $T: V \longrightarrow V'$. To do this, we rewrite the condition (9.1) in

the form $T(v) = \rho_g'^{-1}(T(\rho_g(v))$, or

$$(9.7) \qquad T(v) = g^{-1}(T(gv)).$$

The average is the linear operator $\tilde{T}$ defined by

$$(9.8) \qquad \tilde{T}(v) = \frac{1}{N}\sum_g g^{-1}(T(gv)),$$

where $N = |G|$ as before. If bases for V, V' are given and if the matrices for $\rho_g, \rho_g', T, \tilde{T}$ are $R_g, R_g', A, \tilde{A}$ respectively, then

$$(9.9) \qquad \tilde{A} = \frac{1}{N}\sum_g R_g'^{-1} A R_g.$$

Since compositions of linear transformations and sums of linear transformations are again linear, $\tilde{T}$ is a linear transformation. To show that it is G-invariant, we fix an element $h \in G$ and let $g' = gh$. Reindexing as in the proof of Lemma (2.8),

$$h^{-1}\tilde{T}(hv) = \frac{1}{N}\sum_g h^{-1}g^{-1}(T(ghv)) = \frac{1}{N}\sum_{g'} g'^{-1}(T(g'v)) = \tilde{T}(v).$$

Therefore $\tilde{T}(hv) = h\tilde{T}(v)$. Since h is arbitrary, this shows that $\tilde{T}$ is G-invariant. □

It may happen that we end up with the trivial linear transformation, that is, $\tilde{T} = 0$ though T was not zero. In fact, Schur's Lemma tells us that we *must* get $\tilde{T} = 0$ if ρ and ρ' are irreducible but not isomorphic. We will make good use of this seemingly negative fact in the proof of the orthogonality relations.

When $\rho = \rho'$, the average can often be shown to be nonzero by using this proposition.

(9.10) **Proposition.** Let ρ be a representation of a finite group G on a vector space V, and let $T\colon V \longrightarrow V$ be a linear operator. Define $\tilde{T}$ by formula (9.8). Then trace $\tilde{T}$ = trace T. Thus if the trace of T isn't zero, then $\tilde{T}$ is not zero either.

Proof. We compute as in formula (9.9), with $R' = R$. Since trace A = trace $R_g^{-1}AR_g$, the proposition follows. □

Here is a sample calculation. Let $G = C_3 = \{1, x, x^2\}$, and let $\rho = \rho'$ be the regular representation (Section 6) of G, so that $V = \mathbb{C}^3$ and

$$R_x = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Let T be the linear operator whose matrix is

$$B = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Then the matrix of $\tilde{T}$ is

$$\tilde{B} = \frac{1}{3}(IBI + R_x^{-1}BR_x + R_x^{-2}BR_x^2)$$
$$= \frac{1}{3}(B + R_x^2BR_x + R_xBR_x^2) = \frac{1}{3}\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{bmatrix}.$$

Or, let T be the linear operator whose matrix is the permutation matrix corresponding to the transposition $y = (\mathbf{1}\ \mathbf{2})$. The average over the group is a sum of the three transpositions: $(y + x^{-1}yx + x^{-2}yx)/3 = (y + xy + x^2y)/3$. In this case,

$$P = \frac{1}{3}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \tilde{P} = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Note that $\tilde{B}$ and $\tilde{P}$ commute with R_x as claimed [see (9.4)], though the original matrices P and B do not.

We will now prove the orthogonality relations, Theorem (5.9a). We saw in Section 6 that formula (5.10) is a consequence of these relations.

Let χ, χ' be two nonisomorphic irreducible characters, corresponding to representations ρ, ρ' of G on V, V'. Using the rule $\chi'(g^{-1}) = \overline{\chi'(g)}$, we can rewrite the orthogonality $\langle \chi', \chi \rangle = 0$ to be proved as

$$\text{(9.11)} \qquad \frac{1}{N}\sum_g \chi'(g^{-1})\chi(g) = 0.$$

Now Schur's Lemma asserts that every G-invariant linear transformation $V \longrightarrow V'$ is zero. In particular, the linear transformation $\tilde{T}$ which we obtain by averaging any linear transformation T is zero. Taking into account formula (9.9), this proves the following lemma:

(9.12) Lemma. Let R, R' be nonisomorphic irreducible representations of G. Then

$$\sum_g R_{g^{-1}}{}'AR_g = 0$$

for every matrix A of the appropriate shape. □

Let's warm up by checking orthogonality in the case that ρ and ρ' have dimension 1. In this case, $R_{g'}, R_g{}'$ are 1×1 matrices, that is, scalars, and $\chi(g) = R_g$. If we set $A = 1$, then except for the factor $1/N$, (9.12) becomes (9.11), and we are done.

Lemma (9.12) also implies orthogonality in higher dimensions, but only after a small computation. Let us denote the entries of a matrix M by $(M)_{ij}$, as we did in Section 7 of Chapter 4. Then $\chi(g) = \operatorname{trace} R_g = \Sigma_i\, (R_g)_{ii}$. So $\langle \chi', \chi \rangle$ expands to

$$\text{(9.13)} \qquad \langle \chi', \chi \rangle = \frac{1}{N}\sum_g \sum_{i,j} (R_{g^{-1}}{}')_{ii}(R_g)_{jj}.$$

We may reverse the order of summation. So to prove that $\langle \chi', \chi \rangle = 0$, it suffices to show that for all i, j,

$$\sum_g (R_{g^{-1}}')_{ii}(R_g)_{jj} = 0. \tag{9.14}$$

The proof of the following lemma is elementary:

(9.15) **Lemma.** Let M, N be matrices and let $P = Me_{\alpha\beta}N$, where $e_{\alpha\beta}$ is a matrix unit of suitable size. The entries of P are $(P)_{ij} = (M)_{i\alpha}(N)_{\beta j}$. □

We substitute e_{ij} for A in Lemma (9.12) and apply Lemma (9.15), obtaining

$$0 = (0)_{ij} = \sum_g (R_{g^{-1}}' e_{ij} R_g)_{ij} = \sum_g (R_{g^{-1}}')_{ii}(R_g)_{jj},$$

as required. This shows that $\langle \chi', \chi \rangle = 0$ if χ and χ' are characters of nonisomorphic irreducible representations.

Next, suppose that $\chi = \chi'$. We have to show that $\langle \chi, \chi \rangle = 1$. Averaging A as in (9.9) need not give zero now, but according to Schur's Lemma, it gives a scalar matrix:

$$\frac{1}{N}\sum_g R_{g^{-1}} A R_g = \tilde{A} = aI. \tag{9.16}$$

By Proposition (9.10), trace A = trace $\tilde{A}$, and trace $\tilde{A} = da$, where $d = \dim \rho$. So

$$a = \text{trace } A/d. \tag{9.17}$$

We set $A = e_{ij}$ in (9.16) and apply Lemma (9.15) again, obtaining

$$(aI)_{ij} = \frac{1}{N}\sum_g (R_{g^{-1}} A R_g)_{ij} = \frac{1}{N}\sum_g (R_{g^{-1}})_{ii}(R_g)_{jj}, \tag{9.18}$$

where $a = (\text{trace } e_{ij})/d$. The left-hand side of (9.18) is zero if $i \neq j$ and is equal to $1/d$ if $i = j$. This shows that the terms with $i \neq j$ in (9.13) vanish, and that

$$\langle \chi, \chi \rangle = \frac{1}{N}\sum_g \sum_i (R_{g^{-1}})_{ii}(R_g)_{ii} = \sum_i \left[\frac{1}{N}\sum_g (R_{g^{-1}})_{ii}(R_g)_{ii}\right] = \sum_i 1/d = 1.$$

This completes the proof that the irreducible characters $\chi_1, \chi_2, \ldots$ are orthonormal.

We still have to show that the number of irreducible characters is equal to the number of conjugacy classes, or, equivalently, that the irreducible characters span the space $\mathcal{C}$ of class functions. Let the subspace they span be $\mathcal{X}$. Then [Chapter 7 (2.15)] $\mathcal{C} = \mathcal{X} \oplus \mathcal{X}^\perp$. So we must show that $\mathcal{X}^\perp = 0$, or that a class function ϕ which is orthogonal to every character is zero.

Assume a class function ϕ is given. So ϕ is a complex-valued function on G which is constant on conjugacy classes. Let χ be the character of a representation ρ, and consider the linear operator T: $V \longrightarrow V$ defined by

$$T = \frac{1}{N}\sum_g \overline{\phi(g)}\, \rho_g. \tag{9.19}$$

Its trace is

(9.20) $$\text{trace } T = \frac{1}{N}\sum_g \overline{\phi(g)}\,\chi(g) = <\phi,\chi> = 0,$$

because ϕ is orthogonal to χ.

(9.21) **Lemma.** The operator T defined by (9.19) is G-invariant.

Proof. We have to show (9.2) $\rho_h \circ T = T \circ \rho_h$, or $T = \rho_h^{-1} \circ T \circ \rho_h$, for every $h \in G$. Let $g'' = h^{-1}gh$. Then as g runs over the group G, so does g'', and of course $\rho_h^{-1}\rho_g\rho_h = \rho_{g''}$. Also $\phi(g) = \phi(g'')$ because ϕ is a class function. Therefore

$$\rho_h^{-1}T\rho_h = \frac{1}{N}\sum_g \overline{\phi(g)}\,\rho_h^{-1}\rho_g\rho_h = \frac{1}{N}\sum_{g''} \overline{\phi(g'')}\,\rho_{g''} = T,$$

as required. □

Now if ρ is irreducible as well, then Schur's Lemma (9.6b) applies and shows that $T = cI$. Since trace $T = 0$ (9.20), it follows that $T = 0$. Any representation ρ is a direct sum of irreducible representations, and (9.19) is compatible with direct sums. Therefore $T = 0$ in every case.

We apply this to the case that $\rho = \rho^{\text{reg}}$ is the regular representation. The vector space is $V(G)$. We compute $T(1)$, where 1 denotes the identity element of G. By definition of the regular representation, $\rho_g(1) = g$. So

(9.22) $$0 = T(1) = \frac{1}{N}\sum_g \overline{\phi(g)}\rho_g(1) = \frac{1}{N}\sum_g \overline{\phi(g)}g.$$

Since the elements of G are a basis for $V = V(G)$, this shows that $\overline{\phi(g)} = 0$ for all g, hence that $\phi = 0$. □

10. REPRESENTATIONS OF THE GROUP SU_2

Much of what was done in Sections 6 to 9 carries over without change to *continuous* representations of compact groups G, once a translation-invariant (Haar) measure dg has been found. One just replaces summation by an integral over the group. However, there will be infinitely many irreducible representations if G is not finite.

When we speak of a representation ρ of a compact group, we shall always mean a continuous homomorphism to $GL(V)$, where V is a finite-dimensional complex vector space. The character χ of ρ is then a continuous, complex-valued function on G, which is constant on each conjugacy class. (It is a *class function.*)

For example, the identity map is a two-dimensional representation of SU_2. Its character is the usual trace of 2×2 matrices. We will call this the *standard representation* of SU_2. The conjugacy classes in SU_2 are the sets of matrices with given trace $2c$. They correspond to the latitudes $\{x_1 = c\}$ in the 3-sphere SU_2 [Chapter 8 (2.8)]. Because of this, a class function on SU_2 depends only on x_1. So such a func-

tion can be thought of as a continuous function on the interval $[-1, 1]$. In the notation of Chapter 8 (2.5), the character of the standard representation of SU_2 is

$$\chi(P) = \text{trace } P = a + \overline{a} = 2x_1.$$

Let $|G|$ denote the volume of our compact group G with respect to the measure dg:

$$(10.1) \qquad |G| = \int_G 1 \, dg.$$

Then the hermitian form which replaces (5.8) is

$$(10.2) \qquad \langle \chi, \chi' \rangle = \frac{1}{|G|} \int_G \overline{\chi(g)} \chi'(g) \, dg.$$

With this definition, the orthogonality relations carry over. The proofs of the following extensions to compact groups are the same as for finite groups:

(10.3) **Theorem.**

(a) Every finite-dimensional representation of a compact group G is a direct sum of irreducible representations.

(b) *Schur's Lemma:* Let ρ, ρ' be irreducible representations, and let $T: V \longrightarrow V'$ be a G-invariant linear transformation. Then either T is an isomorphism, or else $T = 0$. If $\rho = \rho'$, then T is multiplication by a scalar.

(c) The characters of the irreducible representations are orthogonal with respect to the form (10.2).

(d) If the characters of two representations are equal, then the representations are isomorphic.

(e) A character χ has the property $\langle \chi, \chi \rangle = 1$ if and only if ρ is irreducible.

(f) If G is abelian, then every irreducible representation is one-dimensional. □

However, the other parts of Theorem (5.9) do not carry over directly. The most significant change in the theory is in Section 6. If G is connected, it cannot operate continuously and nontrivially on a finite set, so finite-dimensional representations can not be obtained from actions on sets. In particular, the regular representation is not finite-dimensional. Analytic methods are needed to extend that part of the theory.

Since a Haar measure is easy to find for the groups U_1 and SU_2, we may consider all of (10.3) proved for them.

Representations of the circle group U_1 are easy to describe, but they are fundamental for an understanding of arbitrary compact groups. It will be convenient to use additive and multiplicative notations interchangeably:

$$(10.4) \qquad SO_2(\mathbb{R}) \xrightarrow{\sim} U_1$$

$$(\text{rotation by } \theta) \rightsquigarrow e^{i\theta} = \alpha.$$

(10.5) **Theorem.** The irreducible representations of U_1 are the nth power maps:

$$U_1 \xrightarrow{\text{“}n\text{”}} U_1,$$

sending $\alpha \rightsquigarrow \alpha^n$, or $\theta \rightsquigarrow n\theta$. There is one such representation for every integer n.

Proof. By (10.3f), the irreducible representations are all one-dimensional, and by (3.5), they are conjugate to unitary representations. Since $GL_1 = \mathbb{C}^\times$ is abelian, conjugation is trivial, so a one-dimensional matrix representation is automatically unitary. Hence an irreducible representation of U_1 is a continuous homomorphism from U_1 to itself. We have to show that the only such homomorphisms are the nth power maps.

(10.6) **Lemma.** The continuous homomorphisms $\psi: \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ are multiplication by a scalar: $\psi(x) = cx$, for some $c \in \mathbb{R}$.

Proof. Let $\psi: \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ be a continuous homomorphism. We will show that $\psi(x) = x\psi(1)$ for all x. This will show that ψ is multiplication by $c = \psi(1)$.

Since ψ is a homomorphism, $\psi(nr) = \psi(r + \cdots + r) = n\psi(r)$, for any real number r and any nonnegative integer n. In particular, $\psi(n) = n\psi(1)$. Also, $\psi(-n) = -\psi(n) = -n\psi(1)$. Therefore $\psi(n) = n\psi(1)$ for every integer n. Next we let $r = m/n$ be a rational number. The $n\psi(r) = \psi(nr) = \psi(m) = m\psi(1)$. Dividing by n, we find $\psi(r) = r\psi(1)$ for every rational number r. Since the rationals are dense in $\mathbb{R}$ and ψ is continuous, $\psi(x) = cx$ for all x. □

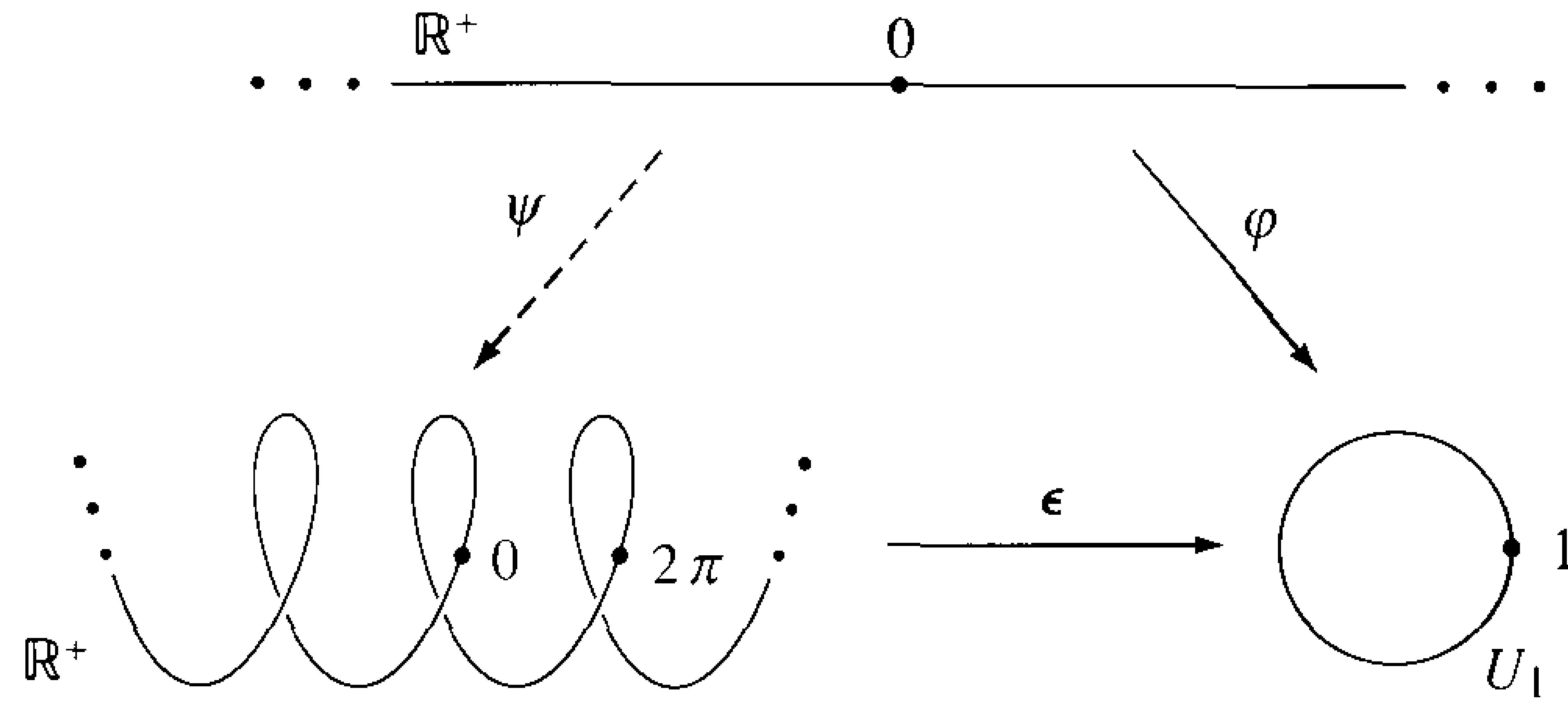
(10.7) **Lemma.** The continuous homomorphisms $\varphi: \mathbb{R}^+ \longrightarrow U_1$ are of the form $\varphi(x) = e^{icx}$ for some $c \in \mathbb{R}$.

Proof. If φ is differentiable, this can be proved using the exponential map of Section 5, Chapter 8. We prove it now for any continuous homomorphism. We consider the exponential homomorphism $\epsilon: \mathbb{R}^+ \longrightarrow U_1$ defined by $\epsilon(x) = e^{ix}$. This homomorphism wraps the real line around the unit circle with period 2π [see Figure (10.8)]. For any continuous function $\varphi: \mathbb{R}^+ \longrightarrow U_1$ such that $\varphi(0) = 1$, there is a unique continuous lifting ψ of this function to the real line such that $\psi(0) = 0$. In other words, we can find a unique continuous function $\psi: \mathbb{R} \longrightarrow \mathbb{R}$ such that $\psi(0) = 0$ and $\varphi(x) = \epsilon(\psi(x))$ for all x. The lifting is constructed starting with the definition $\psi(0) = 0$ and then extending ψ a small interval at a time.

We claim that if φ is a homomorphism, then its lifting ψ is also a homomorphism. If this is shown, then we will conclude that $\psi(x) = cx$ for some c by (10.6), hence that $\varphi(x) = e^{icx}$, as required.

The relation $\varphi(x+y) = \varphi(x)\varphi(y)$ implies that $\epsilon(\psi(x+y)-\psi(x)-\psi(y)) = 1$. Hence $\psi(x + y) - \psi(x)\psi(y) = 2\pi m$ for some integer m which depends continuously on x and y. Varying continuously, m must be constant, and setting $x = y = 0$ shows that $m = 0$. So ψ is a homomorphism, as claimed. □

Now to complete the proof of Theorem (10.5), let $\rho: U_1 \longrightarrow U_1$ be a continuous homomorphism. Then $\varphi = \rho \circ \epsilon: \mathbb{R}^+ \longrightarrow U_1$ is also a continuous homomor-



(10.8) **Figure.**

phism, so $\varphi(x) = e^{icx}$ by (10.7). Moreover, $\varphi(2\pi) = \rho(1)$, which is the case if and only if c is an integer, say n. Then $\rho(e^{ix}) = e^{inx} = (e^{ix})^n$. □

Now let us examine the representations of the group SU_2. Again, there is an infinite family of irreducible representations which arise naturally, and they turn out to form a complete list. Let V_n be the set of homogeneous polynomials of degree n in variables u, v. Such a polynomial will have the form

(10.9) $$f(u,v) = x_0u^n + x_1u^{n-1}v + \cdots + x_nv^n,$$

where the coefficients x_i are complex numbers. Obviously, V_n is a vector space of dimension $n+1$, with basis $(u^n, u^{n-1}v, \ldots, v^n)$. The group $G = GL_2$ operates on V_n in the following way: Let $P \in GL_2$, say

$$P = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Let P act on the basis (u, v) of V_1 as usual:

$$(u', v') = (u, v)P = (au + cv, bu + dv);$$

define ρ_{n_P} by the rule

(10.10) $$u^iv^j \rightsquigarrow u'^iv'^j \quad \text{and}$$

$$f(u,v) \rightsquigarrow x_0u'^n + x_1u'^{n-1}v' + \cdots + x_nv'^n.$$

This is a representation

(10.11) $$\rho_n\colon G \longrightarrow GL(V_n) \approx GL_{n+1}.$$

The trivial representation is ρ_0, and the standard representation is ρ_1.

For example, the matrix of ρ_{2_P} is

(10.12) $$R_{2_P} = \begin{bmatrix} a^2 & ab & b^2 \\ 2ac & ad + bc & 2bd \\ c^2 & cd & d^2 \end{bmatrix}.$$

Its first column is the coordinate vector of $\rho_{2_P}(u^2) = (au + cv)^2 = a^2u^2 + 2acuv + c^2v^2$, and so on.

(10.13) **Theorem.** The representations ρ_n ($n = 0, 1, 2, \ldots$) obtained by restricting (10.11) to the subgroup SU_2 are the irreducible representations of SU_2.

Proof. We consider the subgroup T of SU_2 of diagonal matrices

$$\begin{bmatrix} \alpha & \\ & \overline{\alpha} \end{bmatrix} \tag{10.14}$$

where $\alpha = e^{i\theta}$. This group is isomorphic to U_1. The conjugacy class of an arbitrary unitary matrix P contains two diagonal matrices, namely

$$\begin{bmatrix} \lambda & \\ & \overline{\lambda} \end{bmatrix} \text{ and } \begin{bmatrix} \overline{\lambda} & \\ & \lambda \end{bmatrix},$$

where $\lambda, \overline{\lambda}$ are the eigenvalues of P [Chapter 7 (7.4)]. They coincide only when $\lambda = \pm 1$. So every conjugacy class except $\{I\}$ and $\{-I\}$ intersects T in a pair of matrices.

(10.15) **Proposition.**

(a) A class function on SU_2 is determined by its restriction to the subgroup T.
(b) The restriction of a class function φ to T is an *even* function, which means that

$$\varphi(\alpha) = \varphi(\overline{\alpha}) \quad \text{or} \quad \varphi(\theta) = \varphi(-\theta). \quad \square$$

Next, any representation ρ of SU_2 restricts to a representation on the subgroup T, and T is isomorphic to U_1. The restriction to T of an irreducible representation of SU_2 will usually be reducible, but it can be decomposed into a direct sum of irreducible representations of T. Therefore the restriction of the character χ to T gives us a sum of irreducible characters on U_1. Theorem (10.5) tells us what the irreducible characters of T are: They are the nth powers $e^{in\theta}$, $n \in \mathbb{Z}$. Therefore we find:

(10.16) **Proposition.** The restriction to T of a character χ on SU_2 is a finite sum of exponential functions $e^{in\theta}$. $\square$

Let us calculate the restriction to T of the character χ_n of ρ_n (10.11). The matrix (10.14) acts on monomials by

$$u^i v^j \rightsquigarrow (\alpha^i u^i)(\overline{\alpha}^j v^j) = \alpha^{i-j} u^i v^j.$$

Therefore, its matrix, acting on the basis $(u^n, u^{n-1}v, \ldots, v^n)$, is the diagonal matrix

$$\begin{bmatrix} \alpha^n & & & \\ & \alpha^{n-2} & & \\ & & \ddots & \\ & & & \alpha^{-n} \end{bmatrix},$$

and the value of the character is

$$\chi_n(\alpha) = \alpha^n+\alpha^{n-2}+\cdots+\alpha^{-n} = e^{in\theta}+e^{i(n-2)\theta}+\cdots+e^{-in\theta}, \tag{10.17}$$

or

$$\begin{aligned} \chi_0 &= 1 \\ \chi_1 &= 2\cos\theta = e^{i\theta}+e^{-i\theta} \\ \chi_2 &= 1+2\cos 2\theta = e^{2i\theta}+1+e^{-2i\theta} \\ \chi_3 &= 2\cos 3\theta + 2\cos\theta \\ &\vdots \end{aligned} \tag{10.18}$$

Now let χ' be any irreducible character on SU_2. Its restriction to T is even (10.15b) and is a sum of exponentials $e^{in\theta}$ (10.16). To be even, $e^{in\theta}$ and $e^{-in\theta}$ must occur with the same coefficient, so the character is a linear combination of the functions $\cos n\theta = \frac{1}{2}(e^{in\theta}+e^{-in\theta})$. The functions (10.17) form a basis for the vector space spanned by $\{\cos n\theta\}$. Therefore

$$\chi' = \sum_i r_i\chi_i \tag{10.19}$$

where r_i are rational numbers. A priori, this is true on T, but by (10.15a) it is also true on all of SU_2. Clearing denominators and bringing negative terms to the left in (10.19) yields a relation of the form

$$m\chi' + \sum_j n_j\chi_j = \sum_k n_k\chi_k, \tag{10.20}$$

where n_j, n_k are positive integers and the index sets $\{j\}, \{k\}$ are disjoint. This relation implies

$$m\rho' \oplus \sum_j n_j\rho_j = \sum_k n_k\rho_k.$$

Therefore ρ' is one of the representations ρ_k. This completes the proof of Theorem (10.13). □

We leave the obvious generalizations to the reader.

Israel Herstein

EXERCISES

1. Definition of a Group Representation

1. Let ρ be a representation of a group G. Show that $\det \rho$ is a one-dimensional representation.

2. Suppose that G is a group with a faithful representation by diagonal matrices. Prove that G is abelian.
3. Prove that the rule $S_n \longrightarrow \mathbb{R}^\times$ defined by $p \rightsquigarrow \text{sign}\, p$ is a one-dimensional representation of the symmetric group.
4. Prove that the only one-dimensional representations of the symmetric group S_5 are the trivial representation defined by $\rho(g) = 1$ for all g and the sign representation.
5. (a) Write the standard representation of the octahedral group O by rotations explicitly, choosing a suitable basis for $\mathbb{R}^3$.
   (b) Do the same for the dihedral group D_n.
   *(c) Do the same for the icosahedral group I.
6. Show that the rule $\sigma(\theta) = \begin{bmatrix} \alpha & \alpha^2 - \alpha \\ 0 & \alpha^2 \end{bmatrix}$, $\alpha = e^{i\theta}$, is a representation of SO_2, when a rotation in SO_2 is represented by its angle.
7. Let H be a subgroup of index 2 of a group G, and let $\rho: G \longrightarrow GL(V)$ be a representation. Define $\rho': G \longrightarrow GL(V)$ by the rule $\rho'(g) = \rho(g)$ if $g \in H$, and $\rho'(g) = -\rho(g)$ if $g \notin H$. Prove that ρ' is a representation of G.
8. Prove that every finite group G has a faithful representation on a finite-dimensional complex vector space.
9. Let N be a normal subgroup of a group G. Relate representations of G/N to representations of G.
10. Choose three axes in $\mathbb{R}^3$ passing through the vertices of a regular tetrahedron centered at the origin. (This is not an orthogonal coordinate system.) Find the coordinates of the fourth vertex, and write the matrix representation of the tetrahedral group T in this coordinate system explicitly.

2. G-Invariant Forms and Unitary Representations

1. (a) Verify that the form X^*BY (2.10) is G-invariant.
   (b) Find an orthonormal basis for this form, and determine the matrix P of change of basis. Verify that PAP^{-1} is unitary.
2. Prove the real analogue of (2.2): Let $R: G \longrightarrow GL_n(\mathbb{R})$ be a representation of a finite group G. There is a $P \in GL_n(\mathbb{R})$ such that PR_gP^{-1} is orthogonal for every $g \in G$.
3. Let $\rho: G \longrightarrow SL_2(\mathbb{R})$ be a faithful representation of a finite group by real 2×2 matrices of determinant 1. Prove that G is a cyclic group.
4. Determine all finite groups which have a faithful real two-dimensional representation.
5. Describe the finite groups G which admit faithful real three-dimensional representations with determinant 1.
6. Let V be a hermitian vector space. Prove that the unitary operators on V form a subgroup $U(V)$ of $GL(V)$, and that a representation ρ on V has image in $U(V)$ if and only if the form $\langle\, ,\rangle$ is G-invariant.
7. Let $\langle\, ,\rangle$ be a nondegenerate skew-symmetric form on a vector space V, and let ρ be a representation of a finite group G on V.
   (a) Prove that the averaging process (2.7) produces a G-invariant skew-symmetric form on V.
   (b) Does this prove that every finite subgroup of GL_{2n} is conjugate to a subgroup of SP_{2n}?

8. (a) Let R be the standard two-dimensional representation of D_3, with the triangle situated so that the x-axis is a line of reflection. Rewrite this representation in terms of the basis $x' = x$ and $y' = x + y$.
(b) Use the averaging process to obtain a G-invariant form from dot product in the (x', y')-coordinates.

3. Compact Groups

1. Prove that dx/x is a Haar measure on the multiplicative group $\mathbb{R}^{\times}$.
2. (a) Let $P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$ be a variable 2×2 matrix, and let $dV = dp_{11}dp_{12}dp_{21}dp_{22}$ denote the ordinary volume form on $\mathbb{R}^{2\times 2}$. Show that $(\det P)^{-2}dV$ is a Haar measure on $GL_2(\mathbb{R})$.
(b) Generalize the results of (a).

*3. Show that the form $\dfrac{dx_2 dx_3 dx_4}{x_1}$ on the 3-sphere defines a Haar measure on SU_2. What replaces this expression at points where $x_1 = 0$?

4. Take the complex representation of SO_2 in $\mathbb{R}^2$ given by

$$\sigma(\theta) = \begin{bmatrix} \alpha & \alpha^2 - \alpha \\ 0 & \alpha^2 \end{bmatrix}, \quad \alpha = e^{i\theta},$$

and reduce it to a unitary representation by averaging the hermitian product on $\mathbb{R}^2$.

4. G-Invariant Subspaces and Irreducible Representations

1. Prove that the standard three-dimensional representation of the tetrahedral group T is irreducible as a complex representation.
2. Determine all irreducible representations of a cyclic group C_n.
3. Determine the representations of the icosahedral group I which are not faithful.
4. Let ρ be a representation of a finite group G on a vector space V and let $v \in V$.
(a) Show that averaging gv over G gives a vector $\overline{v} \in V$ which is fixed by G.
(b) What can you say about this vector if ρ is an irreducible representation?
5. Let $H \subset G$ be a subgroup, let ρ be a representation of G on V, and let $v \in V$. Let $w = \Sigma_{h \in H}\, hv$. What can you say about the order of the G-orbit of w?
6. Consider the standard two-dimensional representation of the dihedral group D_n as symmetries of the n-gon. For which values of n is it irreducible as a complex representation?

*7. Let G be the dihedral group D_3, presented as in Chapter 5 (3.6).
(a) Let ρ be an irreducible unitary representation of dimension 2. Show that there is an orthonormal basis of V such that $R_y = \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}$.
(b) Assume that R_y is as above. Use the defining relations $yx = x^2y$, $x^3 = 1$ to determine the possibilities for R_x.
(c) Prove that all irreducible two-dimensional representations of G are isomorphic.
(d) Let ρ be any representation of G, and let $v \in V$ be an eigenvector for the operator ρ_x. Show that v is contained in a G-invariant subspace W of dimension ≤ 2.
(e) Determine all irreducible representations of G.

5. Characters

1. Corollary (5.11) describes a basis for the space of class functions. Give another basis.
2. Find the decomposition of the standard two-dimensional rotation representation of the cyclic group C_n by rotations into irreducible representations.
3. Prove or disprove: Let χ be a character of a finite group G, and define $\overline{\chi}(g) = \overline{\chi(g)}$. Then $\overline{\chi}$ is also a character of G.
4. Find the dimensions of the irreducible representations of the group O of rotations of a cube, the quaternion group, and the dihedral groups D_4, D_5, and D_6.
5. Describe how to produce a unitary matrix by adjusting the entries of a character table.
6. Compare the character tables for the quaternion group and the dihedral group D_4.
7. Determine the character table for D_6.
8. (a) Determine the character table for the groups C_5 and D_5.
   (b) Decompose the restriction of each irreducible character of D_5 into irreducible characters of C_5.
9. (a) Let ρ be a representation of dimension d, with character χ. Prove that the kernel of ρ is the set of group elements such that $\chi(g) = d$.
   (b) Show that if G has a proper normal subgroup, then there is a representation ρ such that ker ρ is a proper subgroup.

*10. Let χ be the character of a representation ρ of dimension d. Prove that $|\chi(g)| \leq d$ for all $g \in G$, and that if $|\chi(g)| = d$, then $\rho(g) = \zeta I$, for some root of unity ζ.

11. Let $G' = G/N$ be a quotient group of a finite group G, and let ρ' be an irreducible representation of G'. Prove that the representation of G defined by ρ' is irreducible in two ways: directly, and using Theorem (5.9).
12. Find the missing rows in the character table below:

	(1) 1	(3) a	(6) b	(6) c	(8) d
χ_1	1	1	1	1	1
χ_2	1	1	-1	-1	1
χ_3	3	-1	1	-1	0
χ_4	3	-1	-1	1	0

13. The table below is a partial character table of a finite group, in which $\zeta = \frac{1}{2}(-1 + \sqrt{3}i)$ and $\gamma = \frac{1}{2}(-1 + \sqrt{7}i)$. The conjugacy classes are all there.

	(1)	(3)	(3)	(7)	(7)
χ_1	1	1	1	ζ	$\overline{\zeta}$
χ_2	3	γ	$\overline{\gamma}$	0	0
χ_3	3	$\overline{\gamma}$	γ	0	0

   (a) Determine the order of the group and the number and the dimensions of the irreducible representations.
   (b) Determine the remaining characters.
   (c) Describe the group by generators and relations.

***14.** Describe the commutator subgroup of a group G in terms of the character table.

***15.** Below is a partial character table. One conjugacy class is missing.

	(1) 1	(1) u	(2) v	(2) w	(3) x
χ_1	1	1	1	1	1
χ_2	1	1	1	1	-1
χ_3	1	-1	1	-1	i
χ_4	1	-1	1	-1	$-i$
χ_5	2	-2	-1	-1	0

(a) Complete the table.
(b) Show that u has order 2, x has order 4, w has order 6, and v has order 3. Determine the orders of the elements in the missing conjugacy class.
(c) Show that v generates a normal subgroup.
(d) Describe the group.

***16.** **(a)** Find the missing rows in the character table below.
(b) Show that the group G with this character table has a subgroup H of order 10, and describe this subgroup as a union of conjugacy classes.
(c) Decide whether H is C_{10} or D_5.
(d) Determine the commutator subgroup of G.
(e) Determine all normal subgroups of G.
(f) Determine the orders of the elements a, b, c, d.
(g) Determine the number of Sylow 2-subgroups and the number of Sylow 5-subgroups of this group.

	(1) 1	(4) a	(5) b	(5) c	(5) d
χ_1	1	1	1	1	1
χ_2	1	1	-1	-1	1
χ_3	1	1	$-i$	i	-1
χ_4	1	1	i	$-i$	-1

***17.** In the character table below, $\zeta = \frac{1}{2}(-1 + \sqrt{3}i)$.

	(1) 1	(6) a	(7) b	(7) c	(7) d	(7) e	(7) f
χ_1	1	1	1	1	1	1	1
χ_2	1	1	1	ζ	$\bar{\zeta}$	ζ	$\bar{\zeta}$
χ_3	1	1	1	$\bar{\zeta}$	ζ	$\bar{\zeta}$	ζ
χ_4	1	1	-1	$-\zeta$	$-\bar{\zeta}$	ζ	$\bar{\zeta}$
χ_5	1	1	-1	$-\bar{\zeta}$	$-\zeta$	$\bar{\zeta}$	ζ
χ_6	1	1	-1	-1	-1	1	1
χ_7	6	-1	0	0	0	0	0

(a) Show that G has a normal subgroup N isomorphic to D_7, and determine the structure of G/N.

(b) Decompose the restrictions of each character to N into irreducible N-characters.
(c) Determine the numbers of Sylow p-subgroups, for $p = 2, 3$, and 7.
(d) Determine the orders of the representative elements c, d, e, f.

6. Permutation Representations and the Regular Representation

1. Verify the values of the characters (6.4) and (6.5).
2. Use the orthogonality relations to decompose the character of the regular representation for the tetrahedral group.
3. Show that the dimension of any irreducible representation of a group G of order $N > 1$ is at most $N - 1$.
4. Determine the character tables for the nonabelian groups of order 12.
5. Decompose the regular representation of C_3 into irreducible *real* representations.
6. Prove Corollary (6.8).
7. Let ρ be the permutation representation associated to the operation of D_3 on itself by conjugation. Decompose the character of ρ into irreducible characters.
8. Let S be a G-set, and let ρ be the permutation representation of G on the space $V(S)$. Prove that the orbit decomposition of S induces a direct sum decomposition of ρ.
9. Show that the standard representation of the symmetric group S_n by permutation matrices is the sum of a trivial representation and an irreducible representation.

*10. Let H be a subgroup of a finite group G. Given an irreducible representation ρ of G, we may decompose its restriction to H into irreducible H-representations. Show that every irreducible representation of H can be obtained in this way.

7. The Representations of the Icosahedral Group

1. Compute the characters χ_2, χ_4, χ_5 of I, and use the orthogonality relations to determine the remaining character χ_3.
2. Decompose the representations of the icosahedral group on the sets of faces, edges, and vertices into irreducible representations.
3. The group S_5 operates by conjugation on its subgroup A_5. How does this action operate on the set of irreducible representations of A_5?

*4. Derive an algorithm for checking that a group is simple by looking at its character table.

5. Use the character table of the icosahedral group to prove that it is a simple group.
6. Let H be a subgroup of index 2 of a group G, and let $\sigma: H \longrightarrow GL(V)$ be a representation. Let a be an element of G not in H. Define a *conjugate* representation $\sigma': H \longrightarrow GL(V)$ by the rule $\sigma'(h) = \sigma(a^{-1}ha)$.
   (a) Prove that σ' is a representation of H.
   (b) Prove that if σ is the restriction to H of a representation of G, then σ' is isomorphic to σ.
   (c) Prove that if b is another element of G not in H, then the representation $\sigma''(h) = \sigma(b^{-1}hb)$ is isomorphic to σ'.
7. (a) Choose coordinates and write the standard three-dimensional matrix representation of the octahedral group O explicitly.

(b) Identify the five conjugacy classes in O, and find the orders of its irreducible representations.

(c) The group O operates on these sets:

(i) six faces of the cube
(ii) three pairs of opposite faces
(iii) eight vertices
(iv) four pairs of opposite vertices
(v) six pairs of opposite edges
(vi) two inscribed tetrahedra

Identify the irreducible representations of O as summands of these representations, and compute the character table for O. Verify the orthogonality relations.

(d) Decompose each of the representations (c) into irreducible representations.

(e) Use the character table to find all normal subgroups of O.

8. **(a)** The icosahedral group I contains a subgroup T, the stabilizer of one of the cubes [Chapter 6 (6.7)]. Decompose the restrictions to T of the irreducible characters of I.

(b) Do the same thing as (a) with a subgroup D_5 of I.

9. Here is the character table for the group $G = PSL_2(\mathbb{F}_7)$, with $\gamma = \frac{1}{2}(-1 + \sqrt{7}i]$, $\gamma' = \frac{1}{2}(-1 - \sqrt{7}i)$.

	(1) 1	(21) a	(24) b	(24) c	(42) d	(56) e
χ_1	1	1	1	1	1	1
χ_2	3	-1	γ	γ'	1	0
χ_3	3	-1	γ'	γ	1	0
χ_4	6	2	-1	-1	0	0
χ_5	7	-1	0	0	-1	1
χ_6	8	0	1	1	0	-1

(a) Use it to give two different proofs that this group is simple.

(b) Identify, so far as possible, the conjugacy classes of the elements

$$\begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}, \begin{bmatrix} 2 & \\ & 4 \end{bmatrix},$$

and find matrices which represent the remaining conjugacy classes.

(c) G operates on the set of one-dimensional subspaces of F^2 ($F = \mathbb{F}_7$). Decompose the associated character into irreducible characters.

8. One-dimensional Representations

1. Prove that the abelian characters of a group G form a group.

2. Determine the character group for the Klein four group and for the quaternion group.

3. Let A,B be matrices such that some power of each matrix is the identity and such that A and B commute. Prove that there is an invertible matrix P such that PAP^{-1} and PBP^{-1} are both diagonal.

4. Let G be a finite abelian group. Show that the order of the character group is equal to the order of G.

*5. Prove that the sign representation $p \rightsquigarrow \text{sign}\, p$ and the trivial representation are the only one-dimensional representations of the symmetric group S_n.

6. Let G be a cyclic group of order n, generated by an element x, and let $\zeta = e^{2\pi i/n}$.
   (a) Prove that the irreducible representations are $\rho_0, \ldots, \rho_{n-1}$, where $\rho_k \colon G \rightsquigarrow \mathbb{C}^\times$ is defined by $\rho_k(x) = \zeta^k$.
   (b) Identify the character group of G.
   (c) Verify the orthogonality relations for G explicitly.

7. (a) Let $\varphi \colon G \longrightarrow G'$ be a homomorphism of abelian groups. Define an induced homomorphism $\hat{\varphi} \colon \hat{G}' \longleftarrow \hat{G}$ between their character groups.
   (b) Prove that $\hat{\varphi}$ is surjective if φ is injective, and conversely.

9. *Schur's Lemma, and Proof of the Orthogonality Relations*

1. Let ρ be a representation of G. Prove or disprove: If the only G-invariant operators on V are multiplication by a scalar, then ρ is irreducible.

2. Let ρ be the standard three-dimensional representation of T, and let ρ' be the permutation representation obtained from the action of T on the four vertices. Prove by averaging that ρ is a summand of ρ'.

3. Let $\rho = \rho'$ be the two-dimensional representation (4.6) of the dihedral group D_3, and let $A = \begin{bmatrix} 1 & 1 \\ & \end{bmatrix}$. Use the averaging process to produce a G-invariant transformation from left multiplication by A.

4. (a) Show that $R_x = \begin{bmatrix} 1 & 1 & -1 \\ & & 1 \\ 1 & & -1 \end{bmatrix}$, $R_y = \begin{bmatrix} & -1 & -1 \\ -1 & & 1 \\ & & -1 \end{bmatrix}$ defines a representation of D_3.
   (b) We may regard the representation ρ_2 of (5.15) as a 1×1 matrix representation. Let T be the linear transformation $\mathbb{C}^1 \longrightarrow \mathbb{C}^3$ whose matrix is $(1, 0, 0)^t$. Use the averaging method to produce a G-invariant linear transformation from T, using ρ_2 and the representation R defined in (a).
   (c) Do part (b), replacing ρ_2 by ρ_1 and ρ_3.
   (d) Decompose R explicitly into irreducible representations.

10. *Representations of the Group SU_2*

1. Determine the irreducible representations of the rotation group SO_3.

2. Determine the irreducible representations of the orthogonal group O_2.

3. Prove that the orthogonal representation $SU_2 \longrightarrow SO_3$ is irreducible, and identify its character in the list (10.18).

4. Prove that the functions (10.18) form a basis for the vector space spanned by $\{\cos n\theta\}$.

5. Left multiplication defines a representation of SU_2 on the space $\mathbb{R}^4$ with coordinates $x_1, \ldots, x_4$, as in Chapter 8, Section 2. Decompose the associated complex representation into irreducible representations.

6. (a) Calculate the four-dimensional volume of the 4-ball of radius r, $B^4 = \{x_1^2 + x_2^2 + x_3^2 + x_4^2 \leq r^2\}$, by slicing with three-dimensional slices.
   (b) Calculate the three-dimensional volume of the 3-sphere S^3, again by slicing. It is

advisable to review the analogous computation of the area of a 2-sphere first. You should find $\frac{d}{dr}$(volume of B^4) = (volume of S^3). If not, try again.

*7. Prove the orthogonality relations for the irreducible characters (10.17) of SU_2 by integration over S^3.

Miscellaneous Problems

*1. Prove that a finite simple group which is not of prime order has no nontrivial representation of dimension 2.

*2. Let H be a subgroup of index 2 of a finite group G, and let a be an element of G not in H, so that aH is the second coset of H in G. Let $S: H \longrightarrow GL_n$ be a matrix representation of H. Define a representation $ind\ S: G \longrightarrow GL_{2n}$ of G, called the *induced* representation, as follows:

$$(ind\ S)_h = \left[\begin{array}{c|c} S_h & \\ \hline & S_{a^{-1}ha} \end{array}\right], \quad (ind\ S)_{ah} = \left[\begin{array}{c|c} & S_{aha} \\ \hline S_h & \end{array}\right].$$

(a) Prove that $ind\ S$ is a representation of G.
(b) Describe the character $\chi_{ind S}$ of $ind\ S$ in terms of the character χ_S of S.
(c) If $R: G \longrightarrow GL_m$ is a representation of G, we may restrict it to H. We denote the restriction by $res\ R: H \longrightarrow GL_n$. Prove that $res\ (ind\ S) \approx S \oplus S'$, where S' is the conjugate representation defined by $S_h' = S_{a^{-1}ha}$.
(d) Prove *Frobenius reciprocity:* $\langle \chi_{ind S}, \chi_R \rangle = \langle \chi_S, \chi_{res R} \rangle$.
(e) Use Frobenius reciprocity to prove that if S and S' are not isomorphic representations, then the induced representation $ind\ S$ of G is irreducible. On the other hand, if $S \approx S'$, then $ind\ S$ is a sum of two irreducible representations R, R'.

*3. Let H be a subgroup of index 2 of a group G, and let R be a matrix representation of G. Let R' denote the *conjugate* representation, defined by $R_g' = R_g$ if $g \in H$, and $R_g' = -R_g$ otherwise.
(a) Show that R' is isomorphic to R if and only if the character of R is identically zero on the coset gH, where $g \notin H$.
(b) Use Frobenius reciprocity to show that $ind(res\ R) \approx R \oplus R'$.
(c) Show that if R is not isomorphic to R', then $res\ R$ is irreducible, and if these two representations are isomorphic, then $res\ R$ is a sum of two irreducible representations of H.

*4. Using Frobenius reciprocity, derive the character table of S_n from that of A_n when (a) $n = 3$, (b) $n = 4$, (c) $n = 5$.

*5. Determine the characters of the dihedral group D_n, using representations induced from C_n.

6. (a) Prove that the only element of SU_2 of order 2 is $-I$.
(b) Consider the homomorphism $\varphi: SU_2 \longrightarrow SO_3$. Let A be an element of SU_2 such that $\varphi(A) = \bar{A}$ has finite order $\bar{n}$ in SO_3. Prove that the order n of A is either $\bar{n}$ or $2\bar{n}$. Also prove that if $\bar{n}$ is even, then $n = 2\bar{n}$.

*7. Let G be a finite subgroup of SU_2, and let $\bar{G} = \varphi(G)$, where $\varphi: SU_2 \longrightarrow SO_3$ is the orthogonal representation (Chapter 8, Section 3). Prove the following.
(a) If $|\bar{G}|$ is even, then $|G| = 2|\bar{G}|$ and $G = \varphi^{-1}(\bar{G})$.

(b) Either $G = \varphi^{-1}(\overline{G})$, or else G is a cyclic group of odd order.

(c) Let G be a cyclic subgroup of SU_2 of order n. Prove that G is conjugate to the group generated by $\begin{bmatrix} \zeta & \\ & \zeta^{-1} \end{bmatrix}$, where $\zeta = e^{2\pi i/n}$.

(d) Show that if $\overline{G}$ is the group D_2, then G is the quaternion group. Determine the matrix representation of the quaternion group H as a subgroup of SU_2 with respect to a suitable orthonormal basis in $\mathbb{C}^2$.

(e) If $\overline{G} = T$, prove that G is a group of order 24 which is not isomorphic to the symmetric group S_4.

***8.** Let ρ be an irreducible representation of a finite group G. How unique is the positive definite G-invariant hermitian form?

***9.** Let G be a finite subgroup of $GL_n(\mathbb{C})$. Prove that if $\Sigma_g \operatorname{tr} g = 0$, then $\Sigma_g\, g = 0$.

***10.** Let $\rho: G \longrightarrow GL(V)$ be a two-dimensional representation of a finite group G, and assume that 1 is an eigenvalue of ρ_g for every $g \in G$. Prove that ρ is a sum of two one-dimensional representations.

***11.** Let $\rho: G \longrightarrow GL_n(\mathbb{C})$ be an irreducible representation of a finite group G. Given any representation $\sigma: GL_n \longrightarrow GL(V)$ of GL_n, we can consider the composition $\sigma \circ \rho$ as a representation of G.

(a) Determine the character of the representation obtained in this way when σ is left multiplication of GL_n on the space $\mathbb{C}^{n\times n}$ of $n \times n$ matrices. Decompose $\sigma \circ \rho$ into irreducible representations in this case.

(b) Find the character of $\sigma \circ \rho$ when σ is the operation of conjugation on $M_n(\mathbb{C})$.

Chapter 10

Rings

Bitte vergiß alles, was Du auf der Schule gelernt hast;
denn Du hast es nicht gelernt.

Edmund Landau

1. DEFINITION OF A RING

The integers form our basic model for the concept of a ring. They are closed under addition, subtraction, and multiplication, but not under division.

Before going to the abstract definition of a ring, we can get some examples by considering subrings of the complex numbers. A *subring* of $\mathbb{C}$ is a subset which is closed under addition, subtraction, and multiplication and which contains 1. Thus any subfield [Chapter 3 (2.1)] is a subring. Another example is the ring of *Gauss integers,* which are complex numbers of the form $a + bi$, where a and b are integers. This ring is denoted by

$$\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}. \tag{1.1}$$

The Gauss integers are the points of a square lattice in the complex plane.

We can form a subring $\mathbb{Z}[\alpha]$ analogous to the ring of Gauss integers, starting with any complex number α. We define $\mathbb{Z}[\alpha]$ to be the smallest subring of $\mathbb{C}$ containing α, and we call it the subring *generated by* α. It is not hard to describe this ring. If a ring contains α, then it contains all positive powers of α because it is closed under multiplication. Also, it contains sums and differences of such powers, and it contains 1. Therefore it contains every complex number β which can be expressed as a polynomial in α with integer coefficients:

$$\beta = a_n\alpha^n + \cdots + a_1\alpha + a_0, \quad \text{where } a_i \in \mathbb{Z}. \tag{1.2}$$

On the other hand, the set of all such numbers is closed under the operations of addition, subtraction, and multiplication, and it contains 1. So it is the subring generated

by α. But $\mathbb{Z}[\alpha]$ will not be represented as a lattice in the complex plane in most cases. For example, the ring $\mathbb{Z}[\frac{1}{2}]$ consists of the rational numbers which can be expressed as a polynomial in $\frac{1}{2}$ with integer coefficients. These rational numbers can be described simply as those whose denominator is a power of 2. They form a dense subset of the real line.

A complex number α is called *algebraic* if it is a root of a polynomial with integer coefficients, that is, if some expression of the form (1.2) is zero. For example, $i + 3, 1/7, 7 + \sqrt[3]{2}$, and $\sqrt{3} + \sqrt{-5}$ are algebraic numbers.

If there is *no* polynomial with integer coefficients having α as a root, then α is called a *transcendental* number. The numbers e and π are transcendental, though it is not easy to prove that they are. If α is transcendental, then two distinct polynomial expressions (1.2) must represent different complex numbers. In this case the elements of the ring $\mathbb{Z}[\alpha]$ correspond bijectively to polynomials $p(x)$ with integer coefficients, by the rule $p(x) \longleftrightarrow p(\alpha)$.

When α is algebraic there will be many polynomial expressions (1.2) which represent the same complex number. For example, when $\alpha = i$, the powers α^n take the four values ± 1, $\pm i$. Using the relation $i^2 = -1$, every expression (1.2) can be reduced to one whose degree in i is ≤ 1. This agrees with the description given above for the ring of Gauss integers.

The two kinds of numbers, algebraic and transcendental, are somewhat analogous to the two possibilities, finite and infinite, for a cyclic group [Chapter 2 (2.7)].

The definition of abstract ring is similar to that of field [Chapter 3 (2.3)], except that multiplicative inverses are not required to exist:

(1.3) **Definition.** A *ring* R is a set with two laws of composition $+$ and $\times$, called addition and multiplication, which satisfy these axioms:

(a) With the law of composition $+$, R is an abelian group, with identity denoted by 0. This abelian group is denoted by R^+.

(b) Multiplication is associative and has an identity denoted by 1.

(c) *Distributive laws:* For all $a, b, c, \in R$,

$$(a + b)c = ac + bc \quad \text{and} \quad c(a + b) = ca + cb.$$

A *subring* of a ring is a subset which is closed under the operations of addition, subtraction, and multiplication and which contains the element 1.

The terminology used is not completely standardized. Some people do not require the existence of a multiplicative identity in a ring. We will study *commutative rings* in most of this book, that is, rings satisfying the commutative law $ab = ba$ for multiplication. So let us agree that the word *ring* will mean *commutative ring with identity,* unless we explicitly mention noncommutativity. The two distributive laws (c) are equivalent for commutative rings.

The ring $\mathbb{R}^{n \times n}$ of all $n \times n$ matrices with real entries is an important example of a ring which is not commutative.

Besides subrings of $\mathbb{C}$, the most important rings are polynomial rings. Given any ring R, a polynomial in x with coefficients in R is an expression of the form

(1.4) $$a_n x^n + \cdots + a_1 x + a_0,$$

with $a_i \in R$. The set of these polynomials forms a ring which is usually denoted by $R[x]$. We will discuss polynomial rings in the next section.

Here are some more examples of rings:

(1.5) **Examples.**

(a) Any field is a ring.

(b) The set $\mathcal{R}$ of continuous real-valued functions of a real variable x forms a ring, with addition and multiplication of functions:

$$[f + g](x) = f(x) + g(x) \quad \text{and} \quad [fg](x) = f(x)g(x).$$

(c) The *zero ring* $R = \{0\}$ consists of a single element 0.

In the definition of a *field* [Chapter 3 (2.3)], the multiplicative identity 1 is required to lie in $F^\times = F - \{0\}$. Hence a field has at least two distinct elements, namely 1 and 0. The relation $1 = 0$ has not been ruled out in a ring, but it occurs only once:

(1.6) **Proposition.** Let R be a ring in which $1 = 0$. Then R is the zero ring.

Proof. We first note that $0a = 0$ for any element a of a ring R. The proof is the same as for vector spaces [Chapter 3 (1.6a)]. Assume that $1 = 0$ in R, and let a be any element of R. Then $a = 1a = 0a = 0$. So every element of R is 0, which means that R is the zero ring. □

Though multiplicative inverses are not required to exist in a ring, a particular element may have an inverse, and the inverse is unique if it exists. Elements which have multiplicative inverses are called *units*. For example, the units in the ring of integers are 1 and -1, and the units in the ring $\mathbb{R}[x]$ of real polynomials are the nonzero constant polynomials. Fields are rings which are not the zero ring and in which every nonzero element is a unit.

The identity element 1 of a ring is always a unit, and any reference to "the" unit element in R refers to the identity. This is ambiguous terminology, but it is too late to change it.

2. FORMAL CONSTRUCTION OF INTEGERS AND POLYNOMIALS

We learn that the ring axioms hold for the integers in elementary school. However, let us look again in order to see what is required in order to write down proofs of properties such as the associative and distributive laws. Complete proofs require a fair amount of writing, and we will only make a start here. It is customary to begin

by defining addition and multiplication for positive integers. Negative numbers are introduced later. This means that several cases have to be treated as one goes along, which is boring, or else a clever notation has to be found to avoid such a case analysis. We will content ourselves with a description of the operations on positive integers. Positive integers are also called *natural numbers.*

The set $\mathbb{N}$ of natural numbers is characterized by these properties, called *Peano's axioms:*

(2.1)

(a) The set $\mathbb{N}$ contains a particular element 1.
(b) *Successor function:* There is a map $\sigma\colon \mathbb{N} \longrightarrow \mathbb{N}$ that sends every integer $n \in \mathbb{N}$ to another integer, called the *next integer* or *successor*. This map is injective, and for every $n \in \mathbb{N}$, $\sigma(n) \neq 1$.
(c) *Induction axiom:* Suppose that a subset S of $\mathbb{N}$ has these properties:
  (i) $1 \in S$;
  (ii) if $n \in S$ then $\sigma(n) \in S$.
  Then S contains every natural number: $S = \mathbb{N}$.

The next integer $\sigma(n)$ will turn into $n + 1$ when addition is defined. At this stage the notation $n + 1$ could be confusing. It is better to use a neutral notation, and we will often denote the successor by n' $[= \sigma(n)]$. Note that σ is assumed to be injective, so if m, n are distinct natural numbers, that is, if $m \neq n$, then m', n' are distinct too.

The successor function allows us to use the natural numbers for counting, which is the basis of arithmetic.

Property (c) is the induction property of the integers. Intuitively, it says that the natural numbers are obtained from 1 by repeatedly taking the next integer: $\mathbb{N} = \{1, 1', 1'', \ldots\}$ $(= \{1, 2, 3, \ldots\})$, that is, counting runs through all natural numbers. This property is the formal basis of induction proofs.

Suppose that a statement P_n is to be proved for every positive integer n, and let S be the set of integers n such that P_n is true. To say that P_n is true for every n is the same as saying that $S = \mathbb{N}$. For this set S, the Induction Axiom translates into the usual induction steps:

(2.2) (i) P_1 is true;
  (ii) if P_n is true then $P_{n'}$ is true.

We can also use Peano's axioms to make recursive definitions. The phrase *recursive definition,* or *inductive definition,* refers to the definition of a sequence of objects C_n indexed by the natural numbers in which each object is defined in terms of the preceding one. The function $C_n = x^n$ is an example. A recursive definition of this function is

$$x^1 = x \quad \text{and} \quad x^{n'} = x^n x.$$

The important points are as follows:

(2.3) (i) C_1 is defined;
(ii) a rule is given for determining $C_{n'}$ $(= C_{n+1})$ from C_n.

It is intuitively clear that (2.3) determines the sequence C_n uniquely, though to prove this from Peano's axioms is tricky. A natural approach to proving it would be as follows: Let S be the set of integers n such that (2.3) determines C_k for every $k \le n$. Then (2.3i) shows that $1 \in S$. Also, (2.3ii) shows that if $n \in S$ then $n' \in S$. The Induction Axiom shows that $S = \mathbb{N}$, hence that C_n is uniquely defined for each n. Unfortunately, the relation $\le$ is not included in Peano's axioms, so it must be defined and its properties derived to start. A proof based on this approach is therefore lengthy, so we won't carry one out here.

Given the set of positive integers and the ability to make recursive definitions, we can define addition and multiplication of positive integers as follows:

(2.4) *Addition:* $m + 1 = m'$ and $m + n' = (m + n)'$.
Multiplication: $m \cdot 1 = m$ and $m \cdot n' = m \cdot n + m$.

In these definitions, we take an arbitrary integer m and then define addition and multiplication for that integer m and for every n recursively. In this way, $m + n$ and $m \cdot n$ are defined for all m and n.

The proofs of the associative, commutative, and distributive laws for the integers are exercises in induction which might be called "Peano playing." We will carry out two of the verifications here as samples.

Proof of the associative law for addition. We are to prove that $(a + b) + n = a + (b + n)$ for all $a, b, n \in \mathbb{N}$. We first check the case $n = 1$ for all a, b. Three applications of definition (2.4) give

$$(a + b) + 1 = (a + b)' = a + b' = a + (b + 1).$$

Next, assume the associative law true for a particular value of n and for all a, b. Then we verify it for n' as follows:

$$\begin{aligned}
(a + b) + n' &= (a + b) + (n + 1) && (\textit{definition})\\
&= ((a + b) + n) + 1 && (\textit{case } n = 1)\\
&= (a + (b + n)) + 1 && (\textit{induction hypothesis})\\
&= a + ((b + n) + 1) && (\textit{case } n = 1)\\
&= a + (b + (n + 1)) && (\textit{case } n = 1)\\
&= a + (b + n') && (\textit{definition}). \square
\end{aligned}$$

Proof of the commutative law for multiplication, assuming that the commutative law for addition has been proved. We first prove the following lemma:

(2.5) $$m' \cdot n = m \cdot n + n.$$

The case $n = 1$ is clear: $m' \cdot 1 = m' = m + 1 = m \cdot 1 + 1$. So assume that (2.5) is true for a particular n and for all values of m. We check it for n':

$$\begin{aligned} m' \cdot n' &= m' \cdot n + m' = m' \cdot n + (m + 1) && \textit{(definition)} \\ &= (m \cdot n + n) + (m + 1) && \textit{(induction)} \\ &= (m \cdot n + m) + (n + 1) && \textit{(various laws for addition)} \\ &= m \cdot n' + n' && \textit{(definition)}. \end{aligned}$$

Next, we check that $1 \cdot n = n$ by induction on n. Finally, we show that $m \cdot n = n \cdot m$ by induction on n, knowing that $m \cdot 1 = m = 1 \cdot m$: Assume it true for n. Then $m \cdot n' = m \cdot n + m = n \cdot m + m = n' \cdot m$, as required. □

The proofs of other properties of addition and multiplication follow similar lines.

We now turn to the definition of polynomial rings. We can define the notion of a *polynomial* with coefficients in any ring R to mean a linear combination of powers of the variable:

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0, \tag{2.6}$$

where $a_i \in R$. Such expressions are often called *formal polynomials,* to distinguish them from polynomial functions. Every formal polynomial with real coefficients determines a polynomial function on the real numbers.

The variable x appearing in (2.6) is an arbitrary symbol, and the monomials x^i are considered independent. This means that if

$$g(x) = b_m x^m + b_{m-1} x^{m-1} + \cdots + b_1 x + b_0$$

is another polynomial with coefficients in R, then $f(x)$ and $g(x)$ are equal if and only if $a_i = b_i$ for all $i = 0, 1, 2, \ldots$.

The *degree* of a nonzero polynomial is the largest integer k such that the coefficient a_k of x^k is not zero. (The degree of the zero polynomial is considered indeterminate.) The coefficient of highest degree of a polynomial which is not zero is called its *leading coefficient,* and a *monic* polynomial is one whose leading coefficient is 1.

The possibility that some of the coefficients of a polynomial may be zero creates a nuisance. We have to disregard terms with zero coefficient: $x^2 + 3 = 0x^3 + x^2 + 3$, for example. So the polynomial $f(x)$ has more than one representation (2.6). One way to standardize notation is to list the nonzero coefficients only, that is, to omit from (2.6) all terms $0x^i$. But zero coefficients may be produced in the course of computations, and they will have to be thrown out. Another possibility is to insist that the highest degree coefficient a_n of (2.6) be nonzero and to list all those of lower degree. The same problem arises. Such conventions therefore require a discussion of special cases in the description of the ring structure. This is irritating, because the ambiguity caused by zero coefficients is not an interesting point.

One way around the notational problem is to list the coefficients of *all* monomials, zero or not. This isn't good for computation, but it allows efficient

verification of the ring axioms. So for the purpose of defining the ring operations, we will write a polynomial in the standard form

$$f(x) = a_0 + a_1x + a_1x^2 + \cdots, \tag{2.7}$$

where the coefficients a_i are all in the ring R and *only finitely many of the coefficients are different from zero.* Formally, the polynomial (2.7) is determined by its vector (or sequence) of coefficients a_i:

$$a = (a_0, a_1, \ldots), \tag{2.8}$$

where $a_i \in R$ and all but a finite number of a_i are zero. Every such vector corresponds to a polynomial. In case R is a field, these infinite vectors form the vector space Z with the infinite basis e_i which was defined in Chapter 3 (5.2d). The vector e_i corresponds to the *monomial* x^i, and the monomials form a basis of the space of all polynomials.

Addition and multiplication of polynomials mimic the familiar operations on real polynomial functions. Let $f(x)$ be as above, and let

$$g(x) = b_0 + b_1x + b_2x^2 + \cdots \tag{2.9}$$

be another polynomial with coefficients in the same ring R, determined by the vector $b = (b_0, b_1, \ldots)$. The *sum* of f and g is

$$\begin{aligned} f(x) + g(x) &= (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots \\ &= \sum_k (a_k + b_k)x^k, \end{aligned} \tag{2.10}$$

which corresponds to vector addition: $a + b = (a_0 + b_0, a_1 + b_1, \ldots)$.

The *product* of two polynomials f, g is computed by multiplying term by term and collecting coefficients of the same degree in x. If we expand the product using the distributive law, but without collecting terms, we obtain

$$f(x)g(x) = \sum_{i,j} a_ib_jx^{i+j}. \tag{2.11}$$

Note that there are finitely many nonzero coefficients a_ib_j. This is a correct formula, but the right side is not in the standard form (2.7) because the same monomial x^n appears many times—once for each pair i, j of indices such that $i + j = n$. So terms have to be collected to put the right side back into standard form. This leads to the definition

$$f(x)g(x) = p_0 + p_1x + p_2x^2 + \cdots,$$

where

$$p_k = a_0b_k + a_1b_{k-1} + \cdots + a_kb_0 = \sum_{i+j=k} a_ib_j. \tag{2.12}$$

However, it may be desirable to defer the collection of terms for a while when making computations.

(2.13) **Proposition.** There is a unique commutative ring structure on the set of polynomials $R[x]$ having these properties:

(a) Addition of polynomials is vector addition (2.10).
(b) Multiplication of monomials is given by the rule (2.12).
(c) The ring R is a subring of $R[x]$, when the elements of R are identified with the constant polynomials.

The proof of this proposition is notationally unpleasant without having any interesting features, so we omit it. □

Polynomials are fundamental to the theory of rings, and we must also consider polynomials, such as $x^2y^2 + 4x^3 - 3x^2y - 4y^2 + 2$, in several variables. There is no major change in the definitions.

Let $x_1,\dots,x_n$ be variables. A *monomial* is a formal product of these variables, of the form

$$x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n},$$

where the exponents i_ν are nonnegative integers. The n-tuple $(i_1,\dots i_n)$ of exponents determines the monomial. Such an n-tuple is called a *multi-index*, and vector notation $i = (i_1,\dots,i_n)$ for multi-indices is very convenient. Using it, we may write the monomial symbolically as

$$x^i = x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}. \tag{2.14}$$

The monomial x^0, where $0 = (0,\dots,0)$, is denoted by 1.

A *polynomial* with coefficients in a ring R is a finite linear combination of monomials, with coefficients in R. Using the shorthand notation (2.14), any polynomial $f(x) = f(x_1,\dots,x_n)$ can be written in exactly one way in the form

$$f(x) = \sum_i a_i x^i, \tag{2.15}$$

where i runs through all multi-indices $(i_1,\dots,i_n)$, the coefficients a_i are in R, and only finitely many of these coefficients are different from zero.

A polynomial which is the product of a monomial by a nonzero element of R is also called a *monomial*. Thus

$$m = rx^i \tag{2.17}$$

is a monomial if $r \in R$ is not zero and if x^i is as above (2.14). A monomial can be thought of as a polynomial which has exactly one nonzero coefficient.

Using multi-index notation, formulas (2.10) and (2.12) define addition and multiplication of polynomials in several variables, and the analogue of Proposition (2.13) is true.

The ring of polynomials with coefficients in R is denoted by one of the symbols

$$R[x_1,\dots,x_n] \quad \text{or} \quad R[x], \tag{2.16}$$

where the symbol x is understood to refer to the set of variables $(x_1,\ldots,x_n)$. When no set of variables has been introduced, $R[x]$ refers to the polynomial ring in one variable x.

3. HOMOMORPHISMS AND IDEALS

A *homomorphism* $\varphi: R \longrightarrow R'$ from one ring to another is a map which is compatible with the laws of composition and which carries 1 to 1, that is, a map such that

$$(3.1) \qquad \varphi(a+b) = \varphi(a) + \varphi(b), \quad \varphi(ab) = \varphi(a)\varphi(b), \quad \varphi(1_R) = 1_{R'},$$

for all $a, b \in R$. An *isomorphism* of rings is a bijective homomorphism. If there is an isomorphism $R \longrightarrow R'$, the two rings are said to be *isomorphic*.

A word about the third part of (3.1) is in order. The assumption that a homomorphism φ is compatible with addition implies that it is a group homomorphism $R^+ \longrightarrow R'^+$. We know that a group homomorphism carries the identity to the identity, so $f(0) = 0$. But R is not a group with respect to $\times$, and we can't conclude that $\varphi(1) = 1$ from compatibility with multiplication. So the condition $\varphi(1) = 1$ must be listed separately. For example, the *zero map* $R \longrightarrow R'$ sending all elements of R to zero is compatible with $+$ and $\times$, but it doesn't send 1 to 1 unless $1 = 0$ in R'. The zero map isn't a ring homomorphism unless R' is the zero ring [see (1.6)].

The most important ring homomorphisms are those obtained by evaluating polynomials. Evaluation of real polynomials at a real number a defines a homomorphism

$$(3.2) \qquad \mathbb{R}[x] \longrightarrow \mathbb{R}, \quad \text{sending } p(x) \rightsquigarrow p(a).$$

We can also evaluate real polynomials at a complex number such as i, to obtain a homomorphism

$$(3.3) \qquad \mathbb{R}[x] \longrightarrow \mathbb{C}, \quad \text{sending } p(x) \rightsquigarrow p(i).$$

The general formulation of the principle of evaluation of polynomials is this:

(3.4) **Proposition.** *Substitution Principle:* Let $\varphi: R \longrightarrow R'$ be a ring homomorphism.

(a) Given an element $\alpha \in R'$, there is a unique homomorphism $\Phi: R[x] \longrightarrow R'$ which agrees with the map φ on constant polynomials and which sends $x \rightsquigarrow \alpha$.

(b) More generally, given elements $\alpha_1,\ldots,\alpha_n \in R'$, there is a unique homomorphism $\Phi: R[x_1,\ldots,x_n] \longrightarrow R'$ from the polynomial ring in n variables to R', which agrees with φ on constant polynomials and which sends $x_\nu \rightsquigarrow \alpha_\nu$, for $\nu = 1,\ldots,n$.

Proof. With vector notation for indices, the proof of (b) is the same as that of (a). Let us denote the image of an element $r \in R$ in R' by r'. Using the fact that Φ

is a homomorphism which restricts to φ on R and sends x_ν to α_ν, we find that it acts on a polynomial $f(x) = \Sigma\, r_i x^i$ by sending

$$\sum r_i x^i \rightsquigarrow \sum \varphi(r_i)\alpha^i = \sum r_i{}'\alpha^i. \tag{3.5}$$

In other words, Φ acts on the coefficients of a polynomial as φ, and it substitutes α for x. Since this formula describes Φ for us, we have proved the uniqueness of the substitution homomorphism. To prove its existence, we take this formula as the definition of Φ, and we show that this map is a homomorphism $R[x] \longrightarrow R'$. It is easy to show that Φ sends 1 to 1 and that it is compatible with addition of polynomials. Compatibility with multiplication can be checked using formula (2.11):

$$\Phi(fg) = \Phi(\sum a_i b_j x^{i+j}) = \sum \Phi(a_i b_j x^{i+j}) = \sum_{i,j} a_i{}' b_j{}' \alpha^{i+j}$$

$$= (\sum_i a_i{}'\alpha^i)(\sum_j b_j{}'\alpha^j) = \Phi(f)\Phi(g). \quad \square$$

Here is an example of the Substitution Principle in which the coefficient ring R changes: Let $\psi: R \longrightarrow R_1$ be a ring homomorphism. Composing ψ with the inclusion of R_1 as a subring of $R_1[x]$, we obtain a homomorphism $\varphi: R \longrightarrow R_1[x]$. The Substitution Principle asserts that there is a unique extension of φ to a homomorphism $\Phi: R[x] \longrightarrow R_1[x]$ which sends $x \rightsquigarrow x$. This is the map which operates on the coefficients of a polynomial, leaving the variable x fixed. If we denote $\psi(a)$ by a', then it sends a polynomial $a_n x^n + \cdots + a_1 x + a_0$ to $a_n{}'x^n + \cdots + a_1{}'x + a_0{}'$.

An important case is the homomorphism $\mathbb{Z} \longrightarrow \mathbb{F}_p$, where $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ is the field with p elements. This map extends to a homomorphism

$$\mathbb{Z}[x] \longrightarrow \mathbb{F}_p[x], \text{ sending} \tag{3.6}$$

$$f(x) = a_n x^n + \cdots + a_0 \rightsquigarrow \overline{a}_n x^n + \cdots + \overline{a}_0 = \bar{f}(x),$$

where $\overline{a}_i$ denotes the residue class of a_i modulo p. It is natural to call the polynomial $\bar{f}(x)$ the *residue of $f(x)$ modulo p*.

The Substitution Principle is also an efficient way to prove that various constructions of polynomial rings are equivalent; the isomorphism

$$R[x, y] \approx R[x][y]$$

is a typical example. Here the right side stands for the ring of polynomials in y whose coefficients are polynomials in x. The statement that these rings are isomorphic is a formalization of the procedure of collecting terms of like degree in y in a polynomial $f(x, y)$, to write it as a polynomial in y. For example,

$$x^2y^2 + 4x^3 - 3x^2y - 4y^2 + 2 = (x^2 - 4)y^2 - (3x^2)y + (4x^3 + 2).$$

(3.7) **Corollary.** *Let $x = (x_1, \ldots, x_m)$ and $y = (y_1, \ldots, y_n)$ denote sets of variables. There is a unique isomorphism $R[x, y] \longrightarrow R[x][y]$ which is the identity on R and which sends the variables to themselves.*

Proof. Note that R is a subring of $R[x]$, and that $R[x]$ is a subring of $R[x][y]$. So R is also a subring of $R[x][y]$. Consider the inclusion map $\varphi: R \longrightarrow R[x][y]$. The Substitution Principle (3.4) tells us that there is a unique homomorphism Φ: $R[x,y] \longrightarrow R[x][y]$ which extends this map and sends the variables x_μ, y_ν wherever we like. So we can send the variables to themselves. The map Φ thus constructed is the required isomorphism. We can show that it has an inverse by using the Substitution Principle once more: We note that $R[x]$ is a subring of $R[x,y]$, so we can extend the inclusion map $\psi: R[x] \longrightarrow R[x,y]$ to a map $\Psi: R[x][y] \longrightarrow R[x,y]$ by sending y_j to itself. The composed homomorphism $\Psi\Phi: R[x,y] \longrightarrow R[x,y]$ is the identity on R and on $\{x_\mu, y_\nu\}$. By the uniqueness of the substitution homomorphism, $\Psi\Phi$ is the identity map. Similarly, $\Phi\Psi$ is the identity. This proves that Φ is an isomorphism. □

Since a real polynomial $f(x)$ can be evaluated at a real number, it defines a polynomial function on the real line. The term *polynomial* is often used to refer to a function obtained in this way, and not much danger is involved in doing this, because we can recover the polynomial from its function:

(3.8) **Proposition.** Let $\mathcal{R}$ denote the ring of continuous real-valued functions on $\mathbb{R}^n$. The map $\varphi: \mathbb{R}[x_1, \ldots, x_n] \longrightarrow \mathcal{R}$ sending a polynomial to its associated polynomial function is an injective homomorphism.

Proof. The existence of this homomorphism follows from the Substitution Principle. Let us prove injectivity. It is enough to show that if the function associated to a polynomial $f(x)$ is the zero function, then $f(x)$ is the zero polynomial. Let the associated function be $\tilde{f}(x)$. If $\tilde{f}(x)$ is identically zero, then all its derivatives are zero too. On the other hand, we can differentiate a formal polynomial by using the rule for differentiating polynomial functions. If some coefficient of our polynomial f is not zero, then the constant term of a suitable derivative will be nonzero too. So that derivative will not vanish at the origin. Therefore $\tilde{f}(x)$ can't be the zero function. □

Another important example of a ring homomorphism is the map from the integers to an arbitrary ring:

(3.9) **Proposition.** There is exactly one homomorphism

$$\varphi: \mathbb{Z} \longrightarrow R$$

from the ring of integers to an arbitrary ring R. It is the map defined by $\varphi(n) =$ "n times 1_R" $= 1_R + \cdots + 1_R$ (n times) if $n > 0$, and $\varphi(-n) = -\varphi(n)$.

Sketch of Proof. Let $\varphi: \mathbb{Z} \longrightarrow R$ be a homomorphism. By the definition of homomorphism, $\varphi(1) = 1_R$, and $\varphi(n + 1) = \varphi(n) + \varphi(1)$. So φ is determined on the natural numbers by the recursive definition

$$\varphi(1) = 1 \quad \text{and} \quad \varphi(n') = \varphi(n) + 1,$$

where $'$ denotes the successor function (2.1b). This formula, together with $\varphi(-n) = -\varphi(n)$ if $n > 0$ and $\varphi(0) = 0$, determines φ uniquely. So the above map is the only possible one. To give a formal proof that this map is a homomorphism, we must go back to Peano's axioms. Let us verify that φ is compatible with addition of positive integers. To prove that $\varphi(m + n) = \varphi(m) + \varphi(n)$, we note that this is true when $n = 1$, by the definition of φ. Assume it true for all m and some particular n. Then we prove it for all m and for n':

$$\begin{aligned} \varphi(m + n') &= \varphi((m + n) + 1) && \textit{(properties of addition of integers)} \\ &= \varphi(m + n) + 1 && \textit{(definition of } \varphi) \\ &= \varphi(m) + \varphi(n) + 1 && \textit{(induction hypothesis)} \\ &= \varphi(m) + \varphi(n') && \textit{(definition of } \varphi). \end{aligned}$$

By induction, $\varphi(m + n) = \varphi(m) + \varphi(n)$ for all m and n. We leave the proof of compatibility with multiplication of positive integers as an exercise. □

This proposition allows us to identify the images of the integers in an arbitrary ring R. Thus we can interpret the symbol 3 as the element $1 + 1 + 1$ in R, and we can interpret an integer polynomial such as $3x^2 + 2x$ as an element of the polynomial ring $R[x]$.

We now go back to an arbitrary ring homomorphism $\varphi: R \longrightarrow R'$. The *kernel* of φ is defined in the same way as the kernel of a group homomorphism:

$$\ker \varphi = \{a \in R \mid \varphi(a) = 0\}.$$

As you will recall, the kernel of a group homomorphism is a subgroup, and in addition it is normal [Chapter 2 (4.9)]. Similarly, the kernel of a ring homomorphism is closed under the ring operations of addition and multiplication, and it also has a stronger property than closure under multiplication:

(3.10) *If $a \in \ker \varphi$ and $r \in R$, then $ra \in \ker \varphi$.*

For if $\varphi(a) = 0$, then $\varphi(ra) = \varphi(r)\varphi(a) = \varphi(r)0 = 0$. On the other hand, $\ker \varphi$ does not contain the unit element 1 of R, and so *the kernel is not a subring*, unless it is the whole ring R. (If $1 \in \ker \varphi$, then $r = r1 \in \ker \varphi$ for all $r \in R$.) Moreover, if $\ker \varphi = R$, then φ is the zero map, and by what was said above, R' is the zero ring.

For example, let φ be the homomorphism $\mathbb{R}[x] \longrightarrow \mathbb{R}$ defined by evaluation at the real number 2. Then $\ker \varphi$ is the set of polynomials which have 2 as a root. It can also be described as the set of polynomials divisible by $x - 2$.

The property of the kernel of a ring homomorphism—that it is closed under multiplication by arbitrary elements of the ring—is abstracted in the concept of an *ideal*. An ideal I of a ring R is, by definition, a subset of R with these properties:

(3.11)

(i) I is a subgroup of R^+;
(ii) If $a \in I$ and $r \in R$, then $ra \in I$.

This peculiar term "ideal" is an abbreviation of "ideal element," which was formerly used in number theory. We will see in Chapter 11 how the term arose. Property (ii) implies that an ideal is closed under multiplication, but it is stronger. A good way to think of properties (i) and (ii) together is this equivalent formulation:

(3.12) *I is not empty, and a linear combination $r_1a_1 + \cdots + r_ka_k$ of elements $a_i \in I$ with coefficients $r_i \in R$ is in I.*

In any ring R, the set of multiples of a particular element a, or equivalently, the set of elements divisible by a, forms an ideal called the *principal ideal* generated by a. This ideal will be denoted in one of the following ways:

(3.13) $$(a) = aR = Ra = \{ra \mid r \in R\}.$$

Thus the kernel of the homomorphism $\mathbb{R}[x] \longrightarrow \mathbb{R}$ defined by evaluation at 2 may be denoted by $(x - 2)$ or by $(x - 2)\mathbb{R}[x]$. Actually the notation (a) for a principal ideal, though convenient, is ambiguous because the ring is not mentioned. For instance, $(x - 2)$ may stand for an ideal in $\mathbb{R}[x]$ or in $\mathbb{Z}[x]$, depending on the circumstances. When there are several rings around, a different notation may be preferable.

We may also consider the ideal I *generated by* a set of elements $a_1, \ldots, a_n$ of R, which is defined to be the smallest ideal containing the elements. It can be described as the set of all linear combinations

(3.14) $$r_1a_1 + \cdots + r_na_n,$$

with coefficients r_i in the ring. For if an ideal contains $a_1, \ldots, a_n$, then (3.12) tells us that it contains every linear combination of these elements. On the other hand, the set of linear combinations is closed under addition, subtraction, and multiplication by elements of R. Hence it is the ideal I. This ideal is often denoted by

(3.15) $$(a_1, \ldots, a_n) = \{r_1a_1 + \cdots + r_na_n \mid r_i \in R\}.$$

For example, if R is the ring $\mathbb{Z}[x]$ of integer polynomials, the notation $(2, x)$ stands for the ideal of linear combinations of 2 and x with integer polynomial coefficients. This ideal can also be described as the set of all integer polynomials $f(x)$ whose constant term is divisible by 2. It is the kernel of the homomorphism $\mathbb{Z}[x] \longrightarrow \mathbb{Z}/2\mathbb{Z}$ defined by $f(x) \rightsquigarrow$ (*residue of* $f(0)$ (modulo 2)).

For the rest of this section, we will describe ideals in some simple cases. In any ring R, the set consisting of zero alone is an ideal, called the *zero ideal*. It is obviously a principal ideal, as is the whole ring. Being generated as an ideal by the element 1, R is called the *unit ideal*, often denoted by (1). The unit ideal is the only ideal which contains a unit. An ideal I is said to be *proper* if it is not (0) or (1).

Fields can be characterized by the fact that they have no proper ideals:

(3.16) **Proposition.**

(a) Let F be a field. The only ideals of F are the zero ideal and the unit ideal.
(b) Conversely, if a ring R has exactly two ideals, then R is a field.

Let us prove (b). Assume that R has exactly two ideals. The properties that distinguish fields among rings are that $1 \neq 0$ and that every nonzero element $a \in R$ has a multiplicative inverse. As we saw above, $1 = 0$ occurs only in the zero ring, which has one element. This ring has only one ideal. Since our ring has two ideals, $1 \neq 0$ in R. The two ideals (1) and (0) are different, so they are the only two ideals of R.

We now show that every nonzero element of R has an inverse. Let $a \in R$ be a nonzero element, and consider the principal ideal (a). Then $(a) \neq (0)$ because $a \in (a)$. Therefore $(a) = (1)$. This implies that 1 is a multiple, say ra, of a. The equation $ar = 1$ shows that a has an inverse. □

(3.17) **Corollary.** Let F be a field and let R' be a nonzero ring. Every homomorphism $\varphi: F \longrightarrow R'$ is injective.

Proof. We apply (3.16). If $\ker \varphi = (1)$, then φ is the zero map. But the zero map isn't a homomorphism because R' isn't the zero ring. Therefore $\ker \varphi = (0)$. □

It is also easy to determine the ideals in the ring of integers.

(3.18) **Proposition.** Every ideal in the ring $\mathbb{Z}$ of integers is a principal ideal.

This is because every subgroup of the additive group $\mathbb{Z}^+$ of integers is of the form $n\mathbb{Z}$ [Chapter 2 (2.3)], and these subgroups are precisely the principal ideals. □

The *characteristic* of a ring R is the nonnegative integer n which generates the kernel of the homomorphism $\varphi: \mathbb{Z} \longrightarrow R$ (3.9). This means that n is the smallest positive integer such that "n times 1_R" $= 0$ or, if the kernel is (0), the characteristic is zero (see Chapter 3, Section 2). Thus $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{Z}$ have characteristic zero, while the field $\mathbb{F}_p$ with p elements has characteristic p.

The proof that every ideal of the ring of integers is principal can be adapted to show that every ideal in the polynomial ring $F[x]$ is principal. To prove this, we need division with remainder for polyomials.

(3.19) **Proposition.** Let R be a ring and let f, g be polynomials in $R[x]$. Assume that the leading coefficient of f is a unit in R. (This is true, for instance, if f is a monic polynomial.) Then there are polynomials $q, r \in R[x]$ such that

$$g(x) = f(x)q(x) + r(x),$$

and such that the degree of the remainder r is less than the degree of f or else $r = 0$.

This division with remainder can be proved by induction on the degree of g. □

Note that when the coefficient ring is a field, the assumption that the leading coefficient of f is a unit is satisfied, provided only that there is a leading coefficient, that is, that $f \neq 0$.

(3.20) **Corollary.** Let $g(x)$ be a monic polynomial in $R[x]$, and let α be an element of R such that $g(\alpha) = 0$. Then $x - \alpha$ divides g in $R[x]$. □

(3.21) **Proposition.** Let F be a field. Every ideal in the ring $F[x]$ of polynomials in a single variable x is a principal ideal.

Proof. Let I be an ideal of $F[x]$. Since the zero ideal is principal, we may assume that $I \neq (0)$. The first step in finding a generator for a nonzero subgroup of $\mathbb{Z}$ is to choose its smallest positive element. Our substitute here is to choose a nonzero polynomial f in I of minimal degree. We claim that I is the principal ideal generated by f. It follows from the definition of an ideal that the principal ideal (f) is contained in I. To prove that $I \subset (f)$, we use division with remainder to write $g = fq + r$, where r has lower degree than f, unless it is zero. Now if g is in the ideal I, then since $f \in I$ the definition of an ideal shows that $r = g - fq$ is in I too. Since f has minimal degree among nonzero elements, the only possibility is that $r = 0$. Thus f divides g, as required. □

The proof of the following corollary is similar to that of (2.6) in Chapter 2.

(3.22) **Corollary.** Let F be a field, and let f,g be polynomials in $F[x]$ which are not both zero. There is a unique monic polymomial $d(x)$ called the *greatest common divisor* of f and g, with the following properties:

(a) d generates the ideal (f, g) of $F[x]$ generated by the two polynomials f, g.
(b) d divides f and g.
(c) If h is any divisor of f and g, then h divides d.
(d) There are polynomials $p,q \in F[x]$ such that $d = pf + qg$. □

4. QUOTIENT RINGS AND RELATIONS IN A RING

Let I be an ideal of a ring R. The cosets of the additive subgroup I^+ of R^+ are the subsets

$$a + I, \quad a \in R.$$

It follows from what has been proved for groups that the set of cosets $R/I = \overline{R}$ is a group under addition. It is also a ring:

(4.1) **Theorem.** Let I be an ideal of a ring R.

(a) There is a unique ring structure on the set of cosets $\overline{R} = R/I$ such that the canonical map $\pi: R \longrightarrow \overline{R}$ sending $a \rightsquigarrow \overline{a} = a + I$ is a homomorphism.
(b) The kernel of π is I.

Proof. This proof has already been carried out in the special case that R is the ring of integers (Chapter 2, Section 9). We want to put a ring structure on $\overline{R}$ with the required properties, and if we forget about multiplication and consider only the addition law, the proof has already been given [Chapter 2 (10.5)]. What is left to do is to define multiplication. Let $x,y \in \overline{R}$, and say that $x = \overline{a} = a + I$ and $y = \overline{b} =$

$b + I$. We would like to define the product to be $xy = \overline{ab} = ab + I$. In contrast with coset multiplication in a group [Chapter 2 (10.1)], the set of products

$$P = \{rs \mid r \in a + I, s \in b + I\}$$

is not always a coset of I. However, as in the case of the ring of integers, the set P is always contained in the single coset $ab + I$: If we write $r = a + u$ and $s = b + v$ with $u, v \in I$, then

$$(a + u)(b + v) = ab + (av + bu + uv),$$

and since I is an ideal, $av + bu + uv \in I$. This is all that is needed to define the product coset: It is the coset which contains the set P. This coset is unique because the cosets partition R. The proof of the remaining assertions closely follows the pattern of Chapter 2, Section 9. □

As in Chapter 6 (8.4) and Chapter 2 (10.9), one can show the following:

(4.2) **Proposition.** *Mapping property of quotient rings:* Let $f: R \longrightarrow R'$ be a ring homomorphism with kernel I and let J be an ideal which is contained in I. Denote the residue ring R/J by $\overline{R}$.

(a) There is a unique homomorphism $\bar{f}: \overline{R} \longrightarrow R'$ such that $\bar{f}\pi = f$:

$$\begin{array}{ccc} R & \xrightarrow{\quad f \quad} & R' \\ & {\scriptstyle \pi} \searrow \quad \nearrow {\scriptstyle \bar{f}} & \\ & \overline{R} = R/J & \end{array}$$

(b) *First Isomorphism Theorem:* If $J = I$, then $\bar{f}$ maps $\overline{R}$ isomorphically to the image of f. □

We will now describe the fundamental relationship between ideals in a quotient ring R/J and ideals in the original ring R.

(4.3) **Proposition.** *Correspondence Theorem:* Let $\overline{R} = R/J$, and let π denote the canonical map $R \longrightarrow \overline{R}$.

(a) There is a bijective correspondence between the set of ideals of R which contain J and the set of all ideals of $\overline{R}$, given by

$$I \rightsquigarrow \pi(I) \quad \text{and} \quad \pi^{-1}(\bar{I}) \leftsquigarrow \bar{I}.$$

(b) If $I \subset R$ corresponds to $\bar{I} \subset \overline{R}$, then R/I and $\overline{R}/\bar{I}$ are isomorphic rings.

The second part of this proposition is often called the *Third Isomorphism Theorem.* [There is also a *Second Isomorphism Theorem* (see Chapter 6, miscellaneous exercise 7)].

Proof. To prove (a), we must check the following points:

(i) If I is an ideal of R which contains J, then $\pi(I)$ is an ideal of $\overline{R}$.
(ii) If $\bar{I}$ is an ideal of $\overline{R}$, then $\pi^{-1}(\bar{I})$ is an ideal of R.
(iii) $\pi^{-1}(\pi(\bar{I})) = I$ and $\pi(\pi^{-1}(\bar{I})) = \bar{I}$.

We know that the image of a subgroup is a subgroup [Chapter 2 (4.4)]. So to show that $\pi(I)$ is an ideal of $\overline{R}$, we need only prove that it is closed under multiplication by elements of $\overline{R}$. Let $\bar{r} \in \overline{R}$, and let $\bar{x} \in \pi(I)$. We write $\bar{r} = \pi(r)$ for some $r \in R$, and $\bar{x} = \pi(x)$ for some $x \in I$. Then $\overline{rx} = \pi(rx)$ and $rx \in I$. So $\overline{rx} \in \pi(I)$. Note that this proof works for all ideals I of R. We do not need the assumption that $I \supset J$ at this point. However, the fact that π is surjective is essential.

Next, we denote the homomorphism $\overline{R} \longrightarrow \overline{R}/\bar{I}$ by φ, and we consider the composed homomorphism $R \xrightarrow{\pi} \overline{R} \xrightarrow{\varphi} \overline{R}/\bar{I}$. Since π and φ are surjective, so is $\varphi \circ \pi$. Moreover, the kernel of $\varphi \circ \pi$ is the set of elements $r \in R$ such that $\pi(r) \in \bar{I} = \ker \varphi$. By definition, this is $\pi^{-1}(\bar{I})$. Therefore $\pi^{-1}(\bar{I})$, being the kernel of a homomorphism, is an ideal of R. This proves (ii). Also, the First Isomorphism Theorem applies to the homomorphism $\varphi \circ \pi$ and shows that $R/\pi^{-1}(\bar{I})$ is isomorphic to $\overline{R}/\bar{I}$. This proves part (b) of the proposition.

It remains to prove (iii); remember that π^{-1} isn't usually a map. The inclusions $\pi^{-1}(\pi(I)) \supset I$ and $\pi(\pi^{-1}(\bar{I})) \subset \bar{I}$ are general properties of any map of sets and for arbitrary subsets. Moreover, the equality $\pi(\pi^{-1}(\bar{I})) = \bar{I}$ holds for any surjective map of sets. We omit the verification of these facts. The final point, that $\pi^{-1}(\pi(I)) \subset I$, is the one which requires that $I \supset J$. Let $x \in \pi^{-1}(\pi(I))$. Then $\pi(x) \in \pi(I)$, so there is an element $y \in I$ such that $\pi(y) = \pi(x)$. Since π is a homomorphism, $\pi(x - y) = 0$ and $x - y \in J = \ker \pi$. Since $y \in I$ and $J \subset I$, this implies that $x \in I$, as required. □

The quotient construction has an important interpretation in terms of *relations* among elements in a ring R. Let us imagine performing a sequence of operations $+, -, \times$ on some elements of R to get a new element a. If the resulting element a is zero, we say that the given elements are related by the equation

$$a = 0. \tag{4.4}$$

For instance, the elements 2, 3, 6 of the ring $\mathbb{Z}$ are related by the equation $2 \times 3 - 6 = 0$.

Now if the element a is not zero, we may ask whether it is possible to modify R in such a way that (4.4) becomes true. We can think of this process as adding a new relation, which will collapse the ring. For example, the relation $3 \times 4 - 5 = 0$ does not hold in $\mathbb{Z}$, because $3 \times 4 - 5 = 7$. But we can impose the relation $7 = 0$ on the integers. Doing so amounts to working modulo 7.

At this point we can forget about the procedure which led us to the particular element a; let it be an arbitrary element of R. Now when we modify R to impose the relation $a = 0$, we want to keep the operations $+$ and $\times$, so we will have to accept some consequences of this relation. For example, $ra = 0$ and $b + a = b$ are the

consequences of multiplying and adding given elements to both sides of $a = 0$. Performing these operations in succession gives us the consequence

$$b + ra = b. \tag{4.5}$$

If we want to set $a = 0$, we must also set $b + ra = b$ for all $b, r \in R$. Theorem (4.1) tells us that this is enough: There are no other consequences of (4.4). To see this, note that if we fix an element b but let r vary, the set $\{b + ra\}$ is the coset $b + (a)$, where $(a) = aR$ is the principal ideal generated by a. Setting $b + ra = b$ for all r is the same as equating the elements of this coset. This is precisely what happens when we pass from R to the quotient ring $\overline{R} = R/(a)$. The elements of $\overline{R}$ are the cosets $\overline{b} = b + (a)$, and the canonical map $\pi: R \longrightarrow \overline{R}$ carries all the elements $b + ra$ in one coset to the same element $\overline{b} = \pi(b)$. So exactly the right amount of collapsing has taken place in $\overline{R}$. Also, $\overline{a} = 0$, because a is an element of the ideal (a), which is the kernel of π. So it is reasonable to view $\overline{R} = R/(a)$ as the ring obtained by introducing the relation $a = 0$ into R.

If our element a was obtained from some other elements by a sequence of ring operations, as we supposed in (4.4), then the fact that π is a homomorphism implies that the same sequence of operations gives 0 in $\overline{R}$. Thus if $uv + w = a$ for some $u, v, w \in R$, then the relation

$$\overline{u}\,\overline{v} + \overline{w} = 0 \tag{4.6}$$

holds in $\overline{R}$. For, since π is a homomorphism, $\overline{u}\,\overline{v} + \overline{w} = \overline{uv + w} = \overline{a} = 0$.

A good example of this construction is the relation $n = 0$ in the ring of integers $\mathbb{Z}$. The resulting ring is $\mathbb{Z}/n\mathbb{Z}$.

More generally, we can introduce any number of relations $a_1 = \dots = a_n = 0$, by taking the ideal I generated by $a_1, \dots, a_n$ (3.15), which is the set of linear combinations $\{r_1a_1 + \cdots + r_na_n \mid r_i \in R\}$. The quotient ring $\overline{R} = R/I$ should be viewed as the ring obtained by introducing the n relations $a_1 = 0, \dots, a_n = 0$ into R. Since $a_i \in I$, the residues $\overline{a}_i$ are zero. Two elements b, b' of R have the same image in $\overline{R}$ if and only if $b' - b \in I$, or $b' = b + r_1a_1 + \cdots + r_na_n$, for some $r_i \in R$. Thus the relations

$$b + r_1a_1 + \cdots + r_na_n = b \tag{4.7}$$

are the only consequences of $a_1 = \cdots = a_n = 0$.

It follows from the Third Isomorphism Theorem (4.3b) that introducing relations one at a time or all together leads to isomorphic results. To be precise, let a, b be elements of a ring R, and let $\overline{R} = R/(a)$ be the result of killing a. Introducing the relation $\overline{b} = 0$ into the ring $\overline{R}$ leads to the quotient ring $\overline{R}/(\overline{b})$, and this ring is isomorphic to the quotient $R/(a, b)$ obtained by killing a and b at the same time, because (a, b) and $(\overline{b})$ are corresponding ideals [see (4.3)].

Note that the more relations we add, the more collapsing takes place in the map $R \longrightarrow \overline{R}$. If we add them carelessly, the worst that can happen is that we may end up with $I = R$ and $\overline{R} = 0$. All relations $a = 0$ become true when we collapse R to the zero ring.

The procedure of introducing relations will lead to a new ring in most cases. That is why it is so important. But in some simple cases the First Isomorphism Theorem can be used to relate the ring obtained to a more familiar one. We will work out two examples to illustrate this.

Let $R = \mathbb{Z}[i]$ be the ring of Gauss integers, and let $\overline{R}$ be obtained by introducing the relation $1 + 3i = 0$. So $\overline{R} = R/I$ where I is the principal ideal generated by $1 + 3i$. We begin by experimenting with the relation, looking for recognizable consequences. Multiplying $-1 = 3i$ on both sides by $-i$, we obtain $i = 3$. So $i = 3$ in $\overline{R}$. On the other hand, $i^2 = -1$ in R, and hence in $\overline{R}$ too. Therefore $3^2 = -1$, or $10 = 0$, in $\overline{R}$. Since $i = 3$ and $10 = 0$ in $\overline{R}$, it is reasonable to guess that $\overline{R}$ is isomorphic to $\mathbb{Z}/(10) = \mathbb{Z}/10\mathbb{Z}$.

(4.8) **Proposition.** The ring $\mathbb{Z}[i]/(1 + 3i)$ is isomorphic to the ring $\mathbb{Z}/10\mathbb{Z}$ of integers modulo 10.

Proof. Having made this guess, we can prove it by analyzing the homomorphism $\varphi: \mathbb{Z} \longrightarrow \overline{R}$ (3.9). By the First Isomorphism Theorem, im $\varphi \approx \mathbb{Z}/(\ker \varphi)$. So if we show that φ is surjective and that $\ker \varphi = 10\mathbb{Z}$, we will have succeeded. Now every element of $\overline{R}$ is the residue of a Gauss integer $a + bi$. Since $i = 3$ in $\overline{R}$, the residue of $a + bi$ is the same as that of the integer $a + 3b$. This shows that φ is surjective. Next, let n be an element of $\ker \varphi$. Using the fact that $\overline{R} = R/I$, we see that n must be in the ideal I, that is, that n is divisible by $1 + 3i$ in the ring of Gauss integers. So we may write $n = (a + bi)(1 + 3i) = (a - 3b) + (3a + b)i$ for some integers a, b. Since n is an integer, $3a + b = 0$, or $b = -3a$. Thus $n = a(1 - 3i)(1 + 3i) = 10a$, and this shows that $\ker \varphi \subset 10\mathbb{Z}$. On the other hand, we already saw that $10 \in \ker \varphi$. So $\ker \varphi = 10\mathbb{Z}$, as required. □

Another possible way to identify the quotient R/I is to find a ring R' and a homomorphism $\varphi: R \longrightarrow R'$ whose kernel is I. To illustrate this, let $\overline{R} = \mathbb{C}[x, y]/(xy)$. Here the fact that xy is a product can be used to find such a map φ.

(4.10) **Proposition.** The ring $\mathbb{C}[x, y]/(xy)$ is isomorphic to the subring of the product ring $\mathbb{C}[x] \times \mathbb{C}[y]$ consisting of the pairs $(p(x), q(y))$ such that $p(0) = q(0)$.

Proof. We can identify the ring $\mathbb{C}[x, y]/(y)$ easily, because the principal ideal (y) is the kernel of the substitution homomorphism $\varphi: \mathbb{C}[x, y] \longrightarrow \mathbb{C}[x]$ sending $y \rightsquigarrow 0$. By the First Isomorphism Theorem, $\mathbb{C}[x, y]/(y) \approx \mathbb{C}[x]$. Similarly, $\mathbb{C}[x, y]/(x) \approx \mathbb{C}[y]$. So it is natural to look at the homomorphism to the product ring $\varphi: \mathbb{C}[x, y] \longrightarrow \mathbb{C}[x] \times \mathbb{C}[y]$, which is defined by $f(x, y) \rightsquigarrow (f(x, 0), f(0, y))$. The kernel of φ is the intersection of the kernels: $\ker \varphi = (y) \cap (x)$. To be in this intersection, a polynomial must be divisible by both y and x. This just means that it is divisible by xy. So $\ker \varphi = (xy)$. By the First Isomorphism Theorem, $\overline{R} = \mathbb{C}[x, y]/(xy)$ is isomorphic to the image of the homomorphism φ. That image is the subring described in the statement of the proposition. □

Aside from the First Isomorphism Theorem, there are no general methods for identifying a quotient ring, because it will usually not be a familiar ring. The ring $\mathbb{C}[x,y]/(y^2 - x^3 + x)$, for example, is fundamentally different from any ring we have seen up to now.

5. ADJUNCTION OF ELEMENTS

In this section we discuss a procedure which is closely related to the introduction of relations, that of adding new elements to a ring. Our model for this procedure is the construction of the complex field, starting from the real numbers. One obtains $\mathbb{C}$ from $\mathbb{R}$ by adjoining i, and the construction is completely formal. That is, the imaginary number i has no properties other than those forced by the relation

$$(5.1) \qquad i^2 = -1.$$

We are now ready to understand the general principle behind this construction. Let us start with an arbitrary ring R, and consider the problem of building a bigger ring containing the elements of R and also containing a new element, which we denote by α. We will probably want α to satisfy some relations such as (5.1), for instance. A ring R' containing R as a subring is called a *ring extension* of R. So we are looking for a suitable extension.

Sometimes the element α may be available in a ring extension R' that we already know. In that case, our solution is the subring of R' generated by R and α. This subring is denoted by $R[\alpha]$. We have already described this ring in Section 1, in the case $R = \mathbb{Z}$ and $R' = \mathbb{C}$. The description is no different in general: $R[\alpha]$ consists of the elements of R' which have polynomial expressions

$$r_n\alpha^n + \cdots + r_1\alpha + r_0$$

with coefficients r_i in R. But as happens when we first construct $\mathbb{C}$ from $\mathbb{R}$, we may not yet know an extension containing α. Then we must construct it abstractly. Actually, we already did this when we constructed the polynomial ring $R[x]$.

Note that the polynomial ring $R[x]$ is an extension of R and that it is generated by R and x. So the notation $R[x]$ agrees with the one introduced above. Moreover, the Substitution Principle (3.4) tells us that the polynomial ring is the *universal solution* to our problem of adjoining a new element, in the following sense: If α is an element of any ring extension R' of R, then there is a unique map $R[x] \longrightarrow R'$ which is the identity on R and which carries x to α. The image of this map will be the subring $R[\alpha]$.

Let us now consider the question of the relations which we want our new element to satisfy. The variable x in the polynomial ring $R[x]$ satisfies no relations except those, such as $0x = 0$, implied by the ring axioms. This is another way to state the universal property of the polynomial ring. We may want some nontrivial relations. But now that we have the ring $R[x]$ in hand we can add relations to it as we like, using the procedure given in Section 4. We introduce relations by using the quotient construction *on the polynomial ring* $R[x]$. The fact that R gets replaced by

$R[x]$ in the construction complicates things notationally, but aside from this notational complication, nothing is different.

For example, we can construct the complex numbers formally by introducing the relation $x^2 + 1 = 0$ into the ring of real polynomials $\mathbb{R}[x] = P$. To do so, we form the quotient ring $\overline{P} = P/(x^2 + 1)$. The residue of x becomes our element i. Note that the relation $\overline{x^2+1} = \overline{x}^2 + \overline{1} = 0$ holds in $\overline{P}$, because the map $\pi: P \longrightarrow \overline{P}$ is a homomorphism and because $x^2 + 1 \in \ker \pi$. And since $\overline{1}$ is the unit element in $\overline{P}$, our standard notation for the unit element drops the bar. So $\overline{P}$ is obtained from $\mathbb{R}$ by adjoining an element $\overline{x}$ satisfying $\overline{x}^2 + 1 = 0$. In other words, $P \approx \mathbb{C}$ as required.

The fact that the quotient $\mathbb{R}[x]/(x^2 + 1)$ is isomorphic to $\mathbb{C}$ also follows from the First Isomorphism Theorem (4.2b): Substitution (3.4) of i for x defines a surjective homomorphism $\varphi: \mathbb{R}[x] \longrightarrow \mathbb{C}$, whose kernel is the set of real polynomials with i as a root. Now if i is a root of a real polynomial $p(x)$, then $-i$ is also a root. Therefore $x - i$ and $x + i$ both divide $p(x)$. The kernel is the set of real polynomials divisible by $(x - i)(x + i) = x^2 + 1$, which is the principal ideal $(x^2 + 1)$. By the First Isomorphism Theorem, $\mathbb{C}$ is isomorphic to $\mathbb{R}[x]/(x^2 + 1)$.

Another simple example of adjunction of an element was used in Section 6 of Chapter 8, where a formal infinitesimal element satisfying

$$\epsilon^2 = 0 \tag{5.2}$$

was introduced to compute tangent vectors. An element of a ring R is called *infinitesimal* or *nilpotent* if some power is zero, and our procedure allows us to adjoin infinitesimals to a ring. Thus the result of adjoining an element ϵ satisfying (5.2) to a ring R is the quotient ring $R' = R[x]/(x^2)$. The residue of x is the infinitesimal element ϵ. In this ring, the relation $\epsilon^2 = 0$ reduces all polynomial expressions in ϵ to degree <2, so the elements of R' have the form $a + b\epsilon$, with $a, b \in R$. But the multiplication rule [Chapter 8 (6.5)] is different from the rule for multiplying complex numbers.

In general, if we want to adjoin an element α satisfying one or more polynomial relations of the form

$$f(\alpha) = c_n\alpha^n + \cdots + c_1\alpha + c_0 = 0 \tag{5.3}$$

to a ring R, the solution is $R' = R[x]/I$, where I is the ideal in $R[x]$ generated by the polynomials $f(x)$. If α denotes the residue $\overline{x}$ of x in R', then

$$0 = \overline{f(x)} = \overline{c}_n\overline{x}^n + \cdots + \overline{c}_0 = \overline{c}_n\alpha^n + \cdots + \overline{c}_0. \tag{5.4}$$

Here $\overline{c}_i$ is the image in R' of the constant polynomial c_i. So α satisfies the relation in R' which corresponds to the relation (5.3) in R. The ring obtained in this way will often be denoted by

$$R[\alpha] = \textit{ring obtained by adjoining } \alpha \textit{ to } R. \tag{5.5}$$

Several elements $\alpha_1, \ldots, \alpha_m$ can be adjoined by repeating this procedure, or by introducing the appropriate relations in the polynomial ring $R[x_1, \ldots, x_m]$ in m variables all at once.

One of the most important cases is that the new element α is required to satisfy a single *monic* equation of degree $n > 0$. Suppose we want the relation $f(x) = 0$, where f is the monic polynomial

(5.6) $$f(x) = x^n + c_{n-1}x^{n-1} + \cdots + c_1x + c_0.$$

It isn't difficult to describe the ring $R[\alpha]$ precisely in this special case.

(5.7) **Proposition.** Let R be a ring, and let $f(x)$ be a monic polynomial of positive degree n, with coefficients in R. Let $R[\alpha]$ denote the ring obtained by adjoining an element satisfying the relation $f(\alpha) = 0$. The elements of $R[\alpha]$ are in bijective correspondence with vectors $(r_0,\ldots,r_{n-1}) \in R^n$. Such a vector corresponds to the linear combination

$$r_0 + r_1\alpha + r_2\alpha^2 + \cdots + r_{n-1}\alpha^{n-1}, \quad \text{with } r_i \in R.$$

This proposition says that the powers $1,\alpha,\alpha^2,\ldots,\alpha^{n-1}$ form a *basis* for $R[\alpha]$ over R. To multiply two such linear combinations in $R[\alpha]$, we use polynomial multiplication and then divide the product by f. The remainder is the linear combination of $1,\alpha,\ldots,\alpha^{n-1}$ which represents the product. So although addition in R' depends only on the degree, multiplication depends strongly on the particular polynomial f.

For example, let R' be the result of adjoining an element α to $\mathbb{Z}$ satisfying the relation $\alpha^3+3\alpha+1 = 0$. So $R' = \mathbb{Z}[x]/(x^3 + 3x + 1)$. The elements of R' are linear combinations $r_0+r_1\alpha+r_2\alpha^2$, where r_i are integers. Addition of two linear combinations is polynomial addition: $(2+\alpha-\alpha^2) + (1+\alpha) = 3+2\alpha-\alpha^2$, for instance. To multiply, we compute the product using polynomial multiplication: $(2+\alpha-\alpha^2)(1+\alpha) = 2+3\alpha-\alpha^3$. Then we divide by $1+3\alpha+\alpha^3$: $2+3\alpha-\alpha^3 = (1+3\alpha+\alpha^3)(-1) + (3+6\alpha)$. Since $1+3\alpha+\alpha^3 = 0$ in R', the remainder $3 + 6\alpha$ is the linear combination which represents the product.

Or let R' be obtained by adjoining an element α to $\mathbb{F}_5$ with the relation $\alpha^2 - 3 = 0$, that is, $R' = \mathbb{F}_5[x]/(x^2 - 3)$. Here α represents a formal square root of 3. The elements of R' are the 25 linear expressions $a + b\alpha$ in α with coefficients $a,b \in \mathbb{F}_5$. This ring is a field. To prove this, we verify that every nonzero element $a + b\alpha$ of R' is invertible. Note that $(a + b\alpha)(a - b\alpha) = a^2 - 3b^2 \in \mathbb{F}_5$. Moreover, the equation $x^2 = 3$ has no solution in $\mathbb{F}_5$, and this implies that $a^2 - 3b^2 \neq 0$. Therefore $a^2 - 3b^2$ is invertible in $\mathbb{F}_5$ and in R'. This shows that $a + b\alpha$ is invertible too. Its inverse is $(a^2 - 3b^2)^{-1}(a - b\alpha)$.

On the other hand, the same procedure applied to $\mathbb{F}_{11}$ does not yield a field. The reason is that $x^2 - 3 = (x + 5)(x - 5)$ in $\mathbb{F}_{11}[x]$. So if α denotes the residue of x in $R' = \mathbb{F}_{11}[x]/(x^2 - 3)$, then $(\alpha + 5)(\alpha - 5) = 0$. This can be explained intuitively by noting that we constructed R' by adjoining a square root of 3 to $\mathbb{F}_{11}$ when that field already contains the two square roots ± 5. At first glance, one might expect to get $\mathbb{F}_{11}$ back by this procedure. But we haven't told α whether to be equal to 5 or to -5. We've only told it that its square is 3. The relation $(\alpha + 5)(\alpha - 5) = 0$ reflects this ambiguity. □

Proof of Proposition (5.7). Since $R[\alpha]$ is a quotient of the polynomial ring $R[x]$, every element in $R[\alpha]$ is the residue of a polynomial. This means that it can be written in the form $g(\alpha)$ for some polynomial $g(x) \in R[x]$. The relation $f(\alpha) = 0$ can be used to replace any polynomial $g(\alpha)$ of degree $\geq n$ by one of lower degree: We perform division with remainder by $f(x)$ on the polynomial $g(x)$, obtaining an expression of the form $g(x) = f(x)q(x) + r(x)$ (3.19). Since $f(\alpha) = 0$, $g(\alpha) = r(\alpha)$. Thus every element β of $R[\alpha]$ can be written as a polynomial in α, of degree $<n$.

We now show that the principal ideal generated by $f(x)$ contains no element of degree $<n$, and therefore that $g(\alpha) \neq 0$ for every nonzero polynomial $g(x)$ of degree $<n$. This will imply that the expression of degree $<n$ for an element β is unique. The principal ideal generated by $f(x)$ is the set of all multiples hf of f. Suppose $h(x) = b_m x^m + \cdots + b_0$, with $b_m \neq 0$. Then the highest-degree term of $h(x)f(x)$ is $b_m x^{m+n}$, and hence hf has degree $m + n \geq n$. This completes the proof of the proposition. □

It is harder to analyze the structure of the ring obtained by adjoining an element which satisfies a nonmonic polynomial relation. One of the simplest and most important cases is obtained by adjoining a multiplicative inverse of an element to a ring. If an element $a \in R$ has an inverse α, then α satisfies the relation

$$a\alpha - 1 = 0. \tag{5.8}$$

So we can adjoin an inverse by forming the quotient ring $R' = R[x]/(ax - 1)$. The residue of x becomes the inverse α of a. This ring has no basis of the type described in Proposition (5.7), but we can compute in it fairly easily because every element of R' has the form $\alpha^k r$, where $r \in R$ and k is a nonnegative integer: Say that $\beta = r_0 + r_1\alpha + \cdots + r_{n-1}\alpha^{n-1}$, with $r_i \in R$. Then since $a\alpha = 1$, we can also write $\beta = \alpha^{n-1}(r_0 a^{n-1} + r_1 a^{n-2} + \cdots + r_{n-1})$.

One interesting example is that R is a polynomial ring itself, say $R = F[t]$, and that we adjoin an inverse to the variable t. Then $R' = F[t,x]/(xt - 1)$. This ring identifies naturally with the ring $F[t,t^{-1}]$ of *Laurent polynomials* in t. A Laurent polynomial is a polynomial in t and t^{-1} of the form

$$f(t) = \sum_{-n}^{n} a_i t^i = a_{-n}t^{-n} + \cdots + a_{-1}t^{-1} + a_0 + a_1 t + \cdots + a_n t^n. \tag{5.9}$$

We leave the construction of this isomorphism as an exercise.

We must now consider a point which we have suppressed in our discussion of adjunction of elements: When we adjoin an element α to a ring R and impose some relations, will our original R be a subring of the ring $R[\alpha]$ which we obtain? We know that R is contained in the polynomial ring $R[x]$, as the subring of constant polynomials. So the restriction of the canonical map $\pi: R[x] \longrightarrow R[x]/I = R[\alpha]$ to constant polynomials gives us a homomorphism $\psi: R \longrightarrow R[\alpha]$, which is the map $r \rightsquigarrow \bar{r}$ considered above. The kernel of the map $\psi: R \longrightarrow R[\alpha] = R[x]/I$ is easy

to determine in principle. It is the set of constant polynomials in the ideal I:

$$\ker \psi = R \cap I. \tag{5.10}$$

It follows from Proposition (5.7) that ψ is injective, and hence that $\ker \psi = 0$, when α is required to satisfy one monic equation. But ψ is not always injective.

For example, we had better not adjoin an inverse of 0 to a ring. From the equation $0\alpha = 1$ we can conclude that $0 = 1$. The zero element is invertible only in the zero ring, so if we insist on adjoining an inverse of 0, we must end up with the zero ring.

More generally, let a, b be two elements of a ring R whose product ab is zero. Then a is not invertible unless $b = 0$. For, if a^{-1} exists in R, then $b = a^{-1}ab = a^{-1}0 = 0$. It follows that if a product ab of two elements of a ring R is zero, then the procedure of adjoining an inverse of a to R must kill b. This can also be seen directly: The ideal of $R[x]$ generated by $ax - 1$ contains $-b(ax - 1) = b$, which shows that the residue of b in the ring $R[x]/(ax - 1)$ is zero.

For example, $\overline{2}\cdot\overline{3} = 0$ in the ring $\mathbb{Z}/(6)$. If we adjoin $\overline{3}^{-1}$ to this ring, we must kill $\overline{2}$. Killing $\overline{2}$ collapses $\mathbb{Z}/(6)$ to $\mathbb{Z}/(2) = \mathbb{F}_2$. Since $\overline{3} = \overline{1}$ is invertible in $\mathbb{F}_2$, no further action is necessary, and $R' = (\mathbb{Z}/(6))[x]/(\overline{3}x - \overline{1}) \approx \mathbb{F}_2$. Again, this can be checked directly. To do so, we note that the ring R' is isomorphic to $\mathbb{Z}[x]/(6, 3x - 1)$, and we analyze the two relations $6 = 0$ and $3x - 1 = 0$. They imply $6x = 0$ and $6x - 2 = 0$; hence $2 = 0$. Then $2x = 0$ too, and combined with $3x - 1 = 0$, this implies $x - 1 = 0$. Hence the ideal $(6, 3x - 1)$ of $\mathbb{Z}[x]$ contains the elements $(2, x - 1)$. On the other hand, 6 and $3x - 1$ are in the ideal $(2, x - 1)$. So the two ideals are equal, and R' is isomorphic to $\mathbb{Z}[x]/(2, x - 1) \approx \mathbb{F}_2$.

An element a of a ring is called a *zero divisor* if there is a nonzero element b such that $ab = 0$. For example, the residue of 3 is a zero divisor in the ring $\mathbb{Z}/(6)$. The term "zero divisor" is traditional, but it has been poorly chosen, because actually every $a \in R$ divides zero: $0 = a0$.

6. INTEGRAL DOMAINS AND FRACTION FIELDS

The difference between rings and fields is that nonzero elements of a ring R do not necessarily have inverses. In this section we discuss the problem of embedding a given ring R as a subring into a field. We saw in the last section that we can not adjoin the inverse of a zero divisor without killing some elements. So a ring which contains zero divisors can not be embedded into a field.

(6.1) **Definition.** An *integral domain* R is a nonzero ring having no zero divisors. In other words, it has the property that if $ab = 0$, then $a = 0$ or $b = 0$, and also $1 \neq 0$ in R.

For example, any subring of a field is an integral domain.

An integral domain satisfies the *cancellation law:*

(6.2) *If* $ab = ac$ *and* $a \neq 0$, *then* $b = c$.

For, from $ab = ac$ we can deduce $a(b - c) = 0$. Then since $a \neq 0$, it follows that $b - c = 0$. □

(6.3) **Proposition.** Let R be an integral domain. Then the polynomial ring $R[x]$ is an integral domain.

(6.4) **Proposition.** An integral domain with finitely many elements is a field.

We leave the proofs of these propositions as exercises. □

(6.5) **Theorem.** Let R be an integral domain. There exists an embedding of R into a field, meaning an injective homomorphism $R \longrightarrow F$, where F is a field.

We could construct the field by adjoining inverses of all nonzero elements of R, using the procedure described in the last section. But in this case it is somewhat simpler to construct F with fractions. Our model is the construction of the rational numbers as fractions of integers, and once the idea of using fractions is put forward, the construction follows the construction of the rational numbers very closely.

Let R be an integral domain. A *fraction* will be a symbol a/b where $a, b \in R$ and $b \neq 0$. Two fractions $a_1/b_1, a_2/b_2$ are called *equivalent*, $a_1/b_1 \approx a_2/b_2$, if

$$a_1b_2 = a_2b_1.$$

Let us check transitivity of this relation—the reflexive and symmetric properties are clear (see Chapter 2, Section 5). Suppose that $a_1/b_1 \approx a_2/b_2$ and also that $a_2/b_2 \approx a_3/b_3$. Then $a_1b_2 = a_2b_1$ and $a_2b_3 = a_3b_2$. Multiply by b_3 and b_1 to obtain

$$a_1b_2b_3 = a_2b_1b_3 = a_3b_2b_1.$$

Cancel b_2 to get $a_3b_1 = a_1b_3$. Thus $a_1/b_1 \approx a_3/b_3$.

The *field of fractions* F of R is the set of equivalence classes of fractions. As we do with rational numbers, we will speak of fractions $a_1/b_1, a_2/b_2$ as equal elements of F if they are equivalent fractions: $a_1/b_1 = a_2/b_2$ in F means $a_1b_2 = a_2b_1$. Addition and multiplication of fractions is defined as in arithmetic:

$$(a/b)(c/d) = ac/bd, \quad a/b + c/d = \frac{ad + bc}{bd}.$$

Here it must be verified that these rules lead to equivalent answers if a/b and c/d are replaced by equivalent fractions. Then the axioms for a field must be verified. All of these verifications are straightforward exercises. □

Notice that R is contained in F, provided that we identify $a \in R$ with the fraction $a/1$ because $a/1 \approx b/1$ only if $a = b$. The map $a \rightsquigarrow a/1$ is the injective homomorphism referred to in the theorem.

As an example, consider the polynomial ring $K[x]$, where K is any field. This is an integral domain, and its fraction field is called the field of *rational functions* in x, with coefficients in K. This field is usually denoted by

$$(6.6) \qquad K(x) = \left\{ \begin{array}{l} \textit{equivalence classes of fractions } f/g, \textit{ where } f, g \\ \textit{are polynomials and } g \textit{ is not the zero polynomial} \end{array} \right\}.$$

If $K = \mathbb{R}$, then evaluation of a rational function $f(x)/g(x)$ defines an actual function on the real line, wherever $g(x) \neq 0$. But as with polynomials, we should distinguish between the formally defined rational functions, which are fractions of polynomials, and the actual functions which they define by evaluation.

The fraction field is a universal solution to the problem of embedding an integral domain into a field. This is shown by the following proposition:

(6.7) **Proposition.** Let R be an integral domain, with field of fractions F, and let $\varphi: R \longrightarrow K$ be any injective homomorphism of R to a field K. Then the rule

$$\Phi(a/b) = \varphi(a)\varphi(b)^{-1}$$

defines the unique extension of φ to a homomorphism $\Phi: F \longrightarrow K$.

Proof. We must check that this extension is well defined. First, since the denominator of a fraction is not allowed to be zero and since φ is injective, $\varphi(b) \neq 0$ for any fraction a/b. Therefore $\varphi(b)$ is invertible in K, and $\varphi(a)\varphi(b)^{-1}$ is an element of K. Next, we check that equivalent fractions have the same image: If $a_2/b_2 \approx a_1/b_1$, then $a_2b_1 = a_1b_2$; hence $\varphi(a_2)\varphi(b_1) = \varphi(a_1)\varphi(b_2)$, and $\Phi(a_2/b_2) = \varphi(a_2)\varphi(b_2)^{-1} = \varphi(a_1)\varphi(b_1)^{-1} = \Phi(a_1/b_1)$, as required. The facts that Φ is a homomorphism and that it is the unique extension of φ follow easily. □

7. MAXIMAL IDEALS

In this section we investigate surjective homomorphisms

$$(7.1) \qquad \varphi: R \longrightarrow F$$

from a ring R to a field F. Given such a homomorphism, the First Isomorphism Theorem tells us that F is isomorphic to $R/\ker \varphi$. Therefore we can recover F and φ, up to isomorphism, from the kernel. To classify such homomorphisms, we must determine the ideals M such that R/M is a field.

By the Correspondence Theorem (4.3), the ideals of $\overline{R} = R/M$ correspond to ideals of R which contain M. Also, fields are characterized by the property of having exactly two ideals (3.16). So if $\overline{R}$ is a field, there are exactly two ideals containing M, namely M and R. Such an ideal is called maximal.

(7.2) **Definition.** An ideal M is *maximal* if $M \neq R$ but M is not contained in any ideals other than M and R.

(7.3) **Corollary.**

(a) An ideal M of a ring R is maximal if and only if $\overline{R} = R/M$ is a field.
(b) The zero ideal of R is maximal if and only if R is a field. □

The next proposition follows from the fact that all ideals of $\mathbb{Z}$ are principal:

(7.4) **Proposition.** The maximal ideals of the ring $\mathbb{Z}$ of integers are the principal ideals generated by prime integers. □

The maximal ideals of the ring $\mathbb{C}[x]$ of complex polynomials in one variable can also be described very simply:

(7.5) **Proposition.** The maximal ideals of the polynomial ring $\mathbb{C}[x]$ are the principal ideals generated by the linear polynomials $x - a$. The ideal M_a generated by $x - a$ is the kernel of the substitution homomorphism $s_a\colon \mathbb{C}[x] \longrightarrow \mathbb{C}$ which sends $f(x) \rightsquigarrow f(a)$. Thus there is a bijective correspondence between maximal ideals M_a and complex numbers a.

Proof. We first show that every maximal ideal is generated by a linear polynomial $x - a$. Let M be maximal. By Proposition (3.21), M is a principal ideal, generated by the monic polynomial $f \in M$ of least degree. Since every complex polynomial of positive degree has a root, f is divisible by some linear polynomial $x - a$. Then f is in the principal ideal $(x - a)$, and hence $M \subset (x - a)$. Since M is maximal, $M = (x - a)$.

Next, we show that the kernel of the substitution homomorphism s_a is generated by $x - a$: To say that a polynomial g is in the kernel of s_a means that a is a root of g, or that $x - a$ divides g. Thus $x - a$ generates ker s_a. Since the image of s_a is a field, this also shows that $(x - a)$ is a maximal ideal. □

The extension of Proposition (7.5) to several variables is one of the most important theorems about polynomial rings.

(7.6) **Theorem.** *Hilbert's Nullstellensatz:* The maximal ideals of the polynomial ring $\mathbb{C}[x_1,\ldots,x_n]$ are in bijective correspondence with points of complex n-dimensional space. A point $a = (a_1,\ldots,a_n)$ in $\mathbb{C}^n$ corresponds to the kernel of the substitution map $s_a\colon \mathbb{C}[x_1,\ldots,x_n] \longrightarrow \mathbb{C}$, which sends $f(x) \rightsquigarrow f(a)$. The kernel M_a of this map is the ideal generated by the linear polynomials

$$x_1 - a_1,\ldots,x_n - a_n.$$

Proof. Let $a \in \mathbb{C}^n$, and let M_a be the kernel of the substitution map s_a. Since s_a is surjective and $\mathbb{C}$ is a field, M_a is a maximal ideal. Next, let us verify that M_a is generated by the linear polynomials, as asserted. To do so, we expand $f(x)$ in powers of $x_1 - a_1,\ldots,x_n - a_n$, writing

$$f(x) = f(a) + \sum_i c_i(x_i - a_i) + \sum_{i,j} c_{ij}(x_i - a_i)(x_j - a_j) + \cdots.$$

You may recognize this as Taylor's expansion: $c_i = \partial f/\partial x_i$, and so on. The existence of such an expansion can be derived algebraically by substituting $x = a + u$ into f, expanding in powers of the variables u, and then substituting $u = x - a$ back into the result. Note that every term on the right side except $f(a)$ is divisible by at least one of the polynomials $(x_i - a_i)$. So if f is in the kernel of s_a, that is, if $f(a) = 0$, then $f(x)$ is in the ideal which these elements generate. This shows that the polynomials $x_i - a_i$ generate M_a.

It is harder to prove that every maximal ideal is of the form M_a for some point $a \in \mathbb{C}^n$. To do so, let M be any maximal ideal, and let K denote the field $\mathbb{C}[x_1,\ldots,x_n]/M$. We consider the restriction of the canonical map (4.1) $\pi: \mathbb{C}[x_1,\ldots,x_n] \longrightarrow K$ to the subring $\mathbb{C}[x_1]$ of polynomials in one variable:

$$\pi_1: \mathbb{C}[x_1] \longrightarrow K.$$

(7.7) **Lemma.** The kernel of π_1 is either zero or else it is a maximal ideal.

Proof. Assume that the kernel is not zero, and let f be a nonzero element in $\ker \pi_1$. Since K is not the zero ring, $\ker \pi_1$ is not the whole ring. So f is not constant, which implies that it is divisible by a linear polynomial, say $f = (x_1 - a_1)g$. Then $\pi_1(x_1 - a_1)\pi_1(g) = \pi_1(f) = 0$ in K. Since K is a field, $\pi_1(x_1 - a_1) = 0$ or $\pi_1(g) = 0$. So one of the two elements $x_1 - a_1$ or g is in $\ker \pi_1$. By induction on the degree of f, $\ker \pi_1$ contains a linear polynomial. Hence it is a maximal ideal (7.5). □

We are going to show that $\ker \pi_1$ is not the zero ideal. It will follow that M contains a linear polynomial of the form $x_1 - a_1$. Since the index 1 can be replaced by any other index, M contains polynomials of the form $x_\nu - a_\nu$ for every $\nu = 1,\ldots,n$. This will show that M is contained in, and hence equal to, the kernel of a substitution map $f(x) \rightsquigarrow f(a)$, as claimed.

So, suppose $\ker \pi_1 = (0)$. Then π_1 maps $\mathbb{C}[x_1]$ isomorphically to its image, which is a subring of K. According to Proposition (6.7), this map can be extended to the field of fractions of $\mathbb{C}[x]$. Hence K contains a field isomorphic to the field of rational functions $\mathbb{C}(x)$ [see (3.17)].

Now the monomials $x^i = x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$ form a basis of $\mathbb{C}[x_1,\ldots,x_n]$ as a vector space over $\mathbb{C}$ (see Section 2). Thus $\mathbb{C}[x_1,\ldots,x_n]$ has a *countable* basis (Appendix, Section 1). Since K is a quotient of $\mathbb{C}[x_1,\ldots,x_n]$, there is a countable family which spans K as vector space over $\mathbb{C}$, namely the residues of the monomials span this field. We will show that there are *uncountably many linearly independent elements* in $\mathbb{C}(x)$. It will follow [Lemma (7.9)] that $\mathbb{C}(x)$ can not be isomorphic to a subspace of K. This contradiction will show $\ker \pi_1 \neq (0)$.

The fact we need is that the elements of the complex field $\mathbb{C}$ do not form a countable set [Appendix (1.7)]. Using this fact, the following two lemmas will finish the proof:

(7.8) **Lemma.** The uncountably many rational functions $(x - \alpha)^{-1}$, $\alpha \in \mathbb{C}$, are linearly independent.

Proof. A rational function f/g defines an actual function by evaluation, at all points of the complex plane at which $g \neq 0$. The rational function $(x - \alpha)^{-1}$ has a *pole* at α, which means that it takes on arbitrarily large values near α. It is bounded near any other point. Consider a linear combination

$$\sum_{i=1}^{n} \frac{c_i}{x - \alpha_i},$$

where $\alpha_1, \dots, \alpha_n$ are distinct complex numbers and where some coefficient, say c_1, is not zero. The first term of this sum is unbounded near α_1, but the others are bounded there. It follows that the linear combination does not define the zero function; hence it is not zero. □

(7.9) **Lemma.** Let V be a vector space which is spanned by a countable family $\{v_1, v_2, \dots\}$ of vectors. Then every set L of linearly independent vectors in V is finite or countably infinite.

Proof. Let L be a linearly independent subset of V, let V_n be the span of the first n vectors $v_1, \dots, v_n$ and let $L_n = L \cap V_n$. Then L_n is a linearly independent set in a finite-dimensional space V_n, hence it is a finite set [Chapter 3 (3.16)]. Moreover, L is the union of all the L_n's. The union of countably many finite sets is finite or countably infinite. □

8. ALGEBRAIC GEOMETRY

To me algebraic geometry is algebra with a kick.

Solomon Lefschetz

Let V be a subset of complex n-space $\mathbb{C}^n$. If V can be defined as the set of common zeros of a finite number of polynomials in n variables, then it is called an *algebraic variety*, or just a *variety* for short. (I don't know the origin of this unattractive term.) For instance, a complex line in $\mathbb{C}^2$ is, by definition, the set of solutions of a linear equation $ax + by + c = 0$. This is a variety. So is a point. The point (a, b) is the set of common zeros of the two polynomials $x - a$ and $y - b$. We have seen a number of other interesting varieties already. The group $SL_2(\mathbb{C})$, for example, being the locus of solutions of the polynomial equation $x_{11}x_{22} - x_{12}x_{21} - 1 = 0$, is a variety in $\mathbb{C}^4$.

Hilbert's Nullstellensatz provides us with an important link between algebra and geometry. It tells us that the maximal ideals in the polynomial ring $\mathbb{C}[x] = \mathbb{C}[x_1, \dots, x_n]$ correspond to points in $\mathbb{C}^n$. This correspondence can also be used to relate algebraic varieties to quotient rings of the polynomial ring.

(8.1) **Theorem.** Let $f_1, \dots, f_r$ be polynomials in $\mathbb{C}[x_1, \dots, x_n]$, and let V be the variety defined by the system of equations $f_1(x) = 0, \dots, f_r(x) = 0$. Let I be the ideal

$(f_1,\dots,f_r)$ generated by the given polynomials. The maximal ideals of the quotient ring $R = \mathbb{C}[x]/I$ are in bijective correspondence with points of V.

Proof. The maximal ideals of R correspond to those maximal ideals of $\mathbb{C}[x]$ which contain I [Correspondence Theorem (4.3)]. And an ideal will contain I if and only if it contains the generators $f_1,\dots,f_r$ of I. On the other hand, the maximal ideal M_a which corresponds to a point $a \in C^n$ is the kernel of the substitution map $f(x) \rightsquigarrow f(a)$. So $f_i \in M_a$ if and only if $f_i(a) = 0$, which means that $a \in V$. □

This theorem shows that the algebraic properties of the ring R are closely connected with the geometry of V. In principle, all properties of the system of polynomial equations

(8.2) $$f_1(x) = \dots = f_r(x) = 0$$

are reflected in the structure of the ring $R = \mathbb{C}[x]/(f_1,\dots,f_r)$. The theory of this relationship is the field of mathematics called algebraic geometry. We won't take the time to go very far into it here. The important thing for us to learn is that geometric properties of the variety provide information about the ring, and conversely.

The simplest question about a set is whether or not it is empty. So we might ask whether it is possible for a ring to have no maximal ideals at all. It turns out that this happens only for the zero ring:

(8.3) **Theorem.** Let R be a ring. Every ideal I of R which is not the unit ideal is contained in a maximal ideal.

(8.4) **Corollary.** The only ring R having no maximal ideals is the zero ring. □

Theorem (8.3) can be proved using the *Axiom of Choice,* or *Zorn's Lemma.* However, for quotients of polynomial rings it is a consequence of the Hilbert Basis Theorem, which we will prove later [Chapter 12 (5.18)]. Rather than enter into a discussion of the Axiom of Choice, we will defer further discussion of the proof to Chapter 12.

If we put Theorems (8.1) and (8.3) together, we obtain another important corollary:

(8.5) **Corollary.** Let $f_1,\dots,f_r$ be polynomials in $\mathbb{C}[x_1,\dots,x_n]$. If the system of equations $f_1 = \dots = f_r = 0$ has no solution in $\mathbb{C}^n$, then 1 is a linear combination

$$1 = \sum g_i f_i$$

of the f_i, with polynomial coefficients.

For, if the system has no solution, then Theorem (8.1) tells us that there is no maximal ideal containing the ideal $I = (f_1,\dots,f_r)$. By Theorem (8.3), I is the unit ideal. □

Most choices of three polynomials f_1, f_2, f_3 in two variables x, y have no common solutions. If follows that we can usually express 1 as a linear combination $1 = p_1f_1 + p_2f_2 + p_3f_3$, where p_i are polynomials. This is not obvious. For instance, the ideal generated by

$$f_1 = x^2 + y^2 - 1, \quad f_2 = x^2 - y + 1, \quad f_3 = xy - 1 \tag{8.6}$$

is the unit ideal. This can be proved by showing that the set of equations $f_1 = f_2 = f_3 = 0$ has no solution in $\mathbb{C}^2$. If we didn't have the Nullstellensatz, it might take us some time to discover that we could write 1 as a linear combination, with polynomial coefficients, of these three polynomials.

The Nullstellensatz has been reformulated in many ways, and actually the one we gave in the last section is not its original form. Here is the original:

(8.7) **Theorem.** *Classical form of the Nullstellensatz:* Let $f_1,\dots, f_r$ and g be polynomials in $\mathbb{C}[x_1,\dots,x_n]$. Let V be the variety of zeros of $f_1,\dots, f_r$, and let I be the ideal generated by these polynomials. If $g = 0$ identically on V, then some power of g is in the ideal I.

Proof. To prove this we study the ring obtained by inverting the polynomial g, by means of the equation $gy = 1$. Assume that g vanishes identically on V. Consider the $r + 1$ polynomials $f_1(x),\dots, f_r(x), g(x)y - 1$ in the variables $x_1,\dots,x_n, y$. The last is the only polynomial which involves the variable y. Notice that these polynomials have no common zero in C^{n+1}. For, if $f_1,\dots, f_r$ vanish at a point $(a_1,\dots,a_n, b) \in C^{n+1}$, then by hypothesis g vanishes too, and hence $gy - 1$ takes the value -1. Corollary (8.5) applies and tells us that the polynomials $f_1,\dots, f_r, gy - 1$ generate the unit ideal in $\mathbb{C}[x, y]$. So we may write

$$1 = \sum_i p_i(x,y)f_i(x,y) + q(x,y)(g(x)y - 1).$$

We substitute $y = 1/g$ into this equation, obtaining

$$1 = \sum_i p_i(x, g^{-1})f_i(x).$$

We now clear denominators in $p_i(x, g^{-1})$, multiplying both sides of the equation by a sufficiently large power of g. This yields the required polynomial expression

$$g(x)^N = \sum_i h_i(x)f_i(x),$$

where $h_i(x) = g(x)^N p_i(x, g^{-1})$. □

It is not easy to get a good feeling for a general algebraic variety in $\mathbb{C}^n$, but the general shape of a variety in $\mathbb{C}^2$ can be described fairly simply.

(8.8) **Proposition.** Two nonzero polynomials $f(x, y)$, $g(x, y)$ in two variables have only finitely many common zeros, unless they have a nonconstant polynomial factor in common.

If the degrees of f and g are m and n respectively, the number of common zeros is bounded by mn. This is known as the *Bezout* bound. For instance, two conics intersect in at most four points. It is somewhat harder to prove the Bezout bound than just the finiteness, and we won't give a proof.

Proof of Proposition (8.8). We assume that f and g have no common nonconstant factor. Let F denote the field of rational functions in x, the field of fractions of the ring $\mathbb{C}[x]$. It is useful to regard f and g as elements of the polynomial ring $F[y]$ in one variable, because we can use the fact that every ideal of $F[y]$ is principal. Let I denote the ideal generated by f, g in $F[y]$. This is a principal ideal, generated by the greatest common divisor h of f and g in $F[y]$ (3.22). If f and g have no common nonconstant factor in $F[y]$, then I is the unit ideal.

Our assumption is that f and g have no common factor in $\mathbb{C}[x, y]$, not that they have no common factor in $F[y]$, so we need to relate these two properties. Factoring polynomials is one of the topics of the next chapter, so we state the fact which we need here and defer the proof (see Chapter 11 (3.9)).

(8.9) **Lemma.** Let $f, g \in \mathbb{C}[x, y]$, and let F be the field of rational functions in x. If f and g have a common factor in $F[y]$ which is not an element of F, then they have a common nonconstant factor in $\mathbb{C}[x, y]$.

We return to the proof of the proposition. Since our two polynomials f, g have no common factor in $\mathbb{C}[x, y]$, they are relatively prime in $F[y]$, so the ideal I they generate in $F[y]$ is the unit ideal. We may therefore write $1 = rf + sg$, where r, s are elements of $F[y]$. Then r, s have denominators which are polynomials in x alone, and we may clear these denominators, multiplying both sides of the equation by a suitable polynomial $p(x)$. This results in an equation of the form

$$p(x) = u(x, y)f(x, y) + v(x, y)g(x, y),$$

where $u, v \in \mathbb{C}[x, y]$. It follows from this equation that a common zero of f and g must also be a zero of p. But p is a polynomial in x alone, and a polynomial in one variable has only finitely many roots. So the variable x takes on only finitely many values at the common zeros of f, g. The same thing is true of the variable y. It follows that the common zeros form a finite set. □

This proposition shows that the most interesting varieties in $\mathbb{C}^2$ are those which are defined as the zeros of a single polynomial $f(x, y)$. These loci are called *algebraic curves,* or *Riemann surfaces,* and their geometry can be quite subtle. A Riemann surface is two-dimensional, so calling it an algebraic curve would seem to be a misnomer. This use of the term *curve* refers to the fact that such a locus can be described analytically by one *complex* parameter, near a point.

A rough description of such a variety, when f is irreducible, follows. (A polynomial is called irreducible if it is not the product of two nonconstant polynomials.)

We regard $f(x,y)$ as a polynomial in y whose coefficients are polynomials in x, say

$$(8.10) \qquad f(x,y) = u_n(x)y^n + \cdots + u_1(x)y + u_0(x),$$

with $u_i(x) \in \mathbb{C}[x]$.

(8.11) **Proposition.** Let $f(x,y)$ be an irreducible polynomial in $\mathbb{C}[x,y]$ which is not a polynomial in x alone, and let S be the locus of zeros of f in $\mathbb{C}^2$. Let n denote the degree of f, as a polynomial in y.

(a) For every value a of the variable x, there are at most n points of S whose x-coordinate is a.
(b) There is a finite set Δ of values of x such that if $a \notin \Delta$ then there are exactly n points of S whose x-coordinate is a.

Proof. Let $a \in \mathbb{C}$, and consider the polynomial $f(a,y)$. The points $(a,b) \in S$ are those such that b is a root of $f(a,y)$. This polynomial is not identically zero, because if it were, then $x - a$ would divide each of the coefficients $u_i(x)$, and hence it would divide f. But f is assumed to be irreducible. Next, the degree of $f(a,y)$ in y is at most n, and so it has at most n roots. It will have fewer than n roots if either

(8.12)

(i) The degree of $f(a,y)$ is less than n, or
(ii) the degree of $f(a,y)$ is n, but this polynomial has a multiple root.

Case (i) occurs when the leading coefficient $u_n(x)$ vanishes at a, that is, when a is a root of $u_n(x)$. Since u_n is a polynomial in x, there are finitely many such values.

Now a complex number b is a multiple root of a polynomial $h(y)$ [meaning that $(y-b)^2$ divides $h(y)$] if and only if it is a root of $h(y)$ and of its derivative $h'(y)$. The proof of this fact is left as an exercise. In our situation, $h(y) = f(a,y)$. The first variable is fixed, so the derivative is the partial derivative with respect to y. Thus case (ii) occurs at points (a,b) which are common zeros of f and $\partial f/\partial y$. Note that f does not divide the partial derivative $\partial f/\partial y$, because the degree of the partial derivative in y is $n - 1$, which is less than the degree of f in y. Since f is assumed to be irreducible, f and $\partial f/\partial y$ have no nonconstant factor in common. Proposition (8.8) tells us that there are finitely many common zeros. □

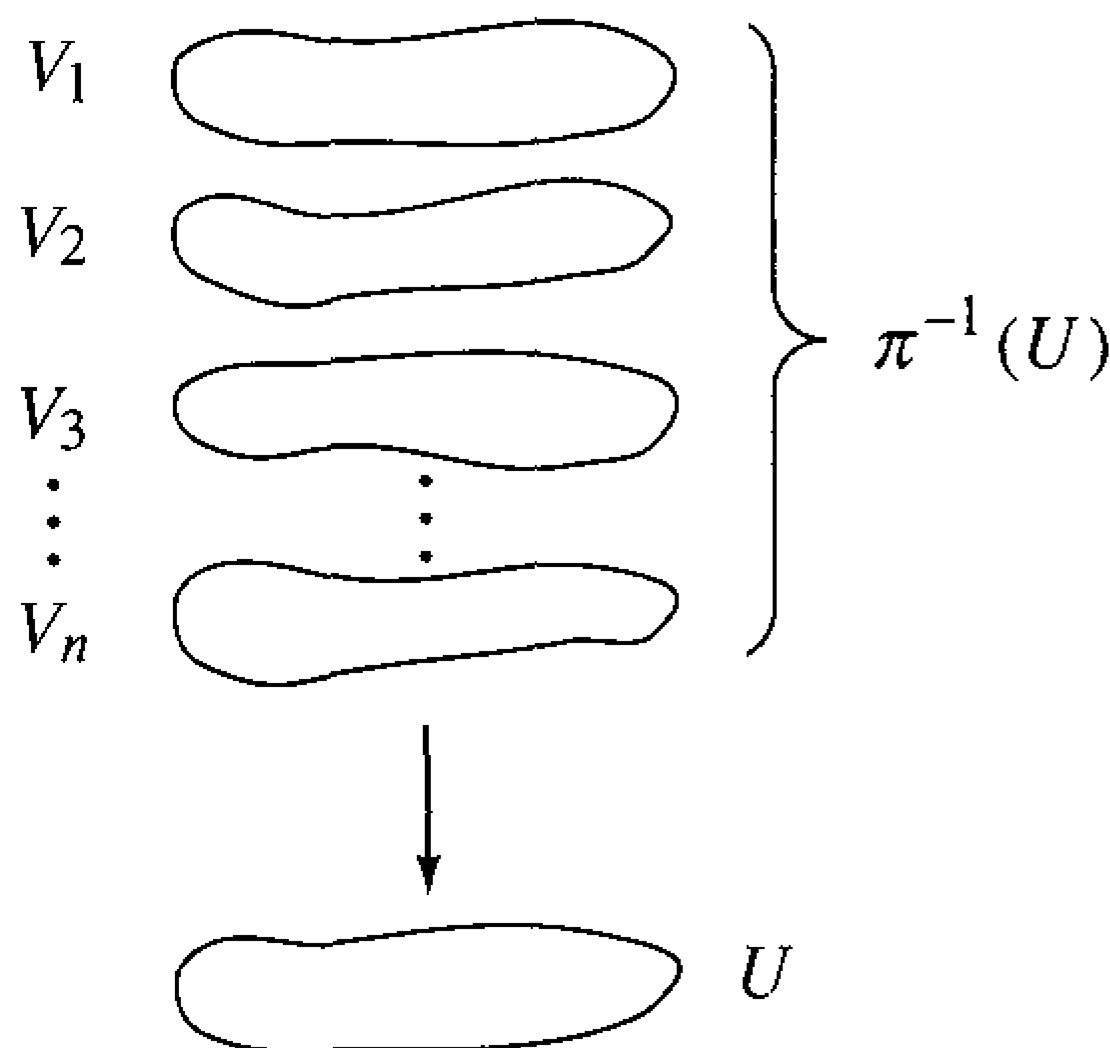
Proposition (8.11) can be summed up by saying that S is an n-sheeted covering of the complex x-plane P. Since there is a finite set Δ above which S has fewer than n sheets, it is called a branched covering. For example, consider the locus $x^2 + xy^2 - 1 = 0$. This equation has two solutions y for every value of x except $x = 0, \pm 1$. There is no solution with $x = 0$, and there is only one with $x = 1$ or -1. So this locus is a branched double covering of P.

Here is the precise definition of a branched covering:

(8.13) **Definition.** An *n-sheeted branched covering* of the complex plane P is a topological space S together with a continuous map $\pi: S \longrightarrow P$, such that

(a) π is n-to-one on the complement of a finite set Δ in P.

(b) For every point $x_0 \in P - \Delta$, there is an open neighborhood U of x, so that $\pi^{-1}(U)$ is made up of n disconnected parts ($\pi^{-1}(U) = V_1 \cup \cdots \cup V_n$), each V_i is open in S, and π maps V_i homeomorphically to U.



(8.14) **Figure.** Part of an n-sheeted covering.

(8.15) **Corollary.** Let $f(x,y)$ be an irreducible polynomial in $\mathbb{C}[x,y]$ which has degree $n > 0$ in the variable y. The Riemann surface of $f(x,y)$ is an n-sheeted branched covering of the plane.

Proof. The fact that the Riemann surface S of f has the first property of a branched covering is Proposition (8.11). So it remains to verify property (8.13b). Consider a point x_0 at which $f(x_0,y)$ has n roots $y_1,\dots,y_n$. Then $(\partial f/\partial y)(x_0,y_1) \neq 0$ because y_1 is not a multiple root of $f(x_0,y_1)$. The Implicit Function Theorem [Appendix (4.1)] applies and tells us that equation (8.2) can be solved for $y = \alpha_1(x)$ as a continuous function of x in some neighborhood U of x_0, in such a way that $y_1 = \alpha_1(x_0)$. Similarly, we can solve for $y = \alpha_i(x)$ such that $y_i = \alpha_i(x_0)$. Cutting down the size of U, we may assume that each $\alpha_i(x)$ is defined on U. Since $y_1,\dots,y_n$ are all distinct and the $\alpha_i(x)$ are continuous functions, they have no common values provided U is made sufficiently small.

Consider the graphs of the n continuous functions α_i:

$$V_i = \{(x,\alpha_i(x)) \mid x \in U\}. \tag{8.16}$$

They are disjoint because the $\alpha_i(x)$ have no common values on U. The map $V_i \longrightarrow U$ is a homeomorphism because it has the continuous inverse function $U \rightsquigarrow V_i$. The inverse sends $x \rightsquigarrow (x,\alpha_i(x))$. And

$$\pi^{-1}(U) = V_1 \cup \cdots \cup V_n$$

because S has at most n points above any x, and the n points have been exhibited as $(x,\alpha_i(x)) \in V_i$. Each of the sets V_i is closed in $U \times \mathbb{C}$, because it is the set of zeros

of the continuous function $y-\alpha_i(x)$. Then V_i is also closed in the subset $\pi^{-1}(U)$ of $U \times \mathbb{C}$. It follows that V_1 is open in $\pi^{-1}(U)$, because it is the complement of the closed set $V_2 \cup \ldots \cup V_n$. Since U is open in $\mathbb{C}$, its inverse image $\pi^{-1}(U)$ is open in S. Thus V_1 is open in an open subset of S, which shows that V_1 is open in S too. Similarly, V_i is open for each i. □

We will look at these loci again in Chapter 13.

In helping geometry, modern algebra is helping itself above all.

Oscar Zariski

EXERCISES

1. Definition of a Ring

1. Prove the following identities in an arbitrary ring R.
   (a) $0a = 0$ (b) $-a = (-1)a$ (c) $(-a)b = -(ab)$
2. Describe explicitly the smallest subring of the complex numbers which contains the real cube root of 2.
3. Let $\alpha = \frac{1}{2}i$. Prove that the elements of $\mathbb{Z}[\alpha]$ form a dense subset of the complex plane.
4. Prove that $7 + \sqrt[3]{2}$ and $\sqrt{3} + \sqrt{-5}$ are algebraic numbers.
5. Prove that for all integers n, $\cos(2\pi/n)$ is an algebraic number.
6. Let $\mathbb{Q}[\alpha,\beta]$ denote the smallest subring of $\mathbb{C}$ containing $\mathbb{Q}$, $\alpha = \sqrt{2}$, and $\beta = \sqrt{3}$, and let $\gamma = \alpha + \beta$. Prove that $\mathbb{Q}[\alpha,\beta] = \mathbb{Q}[\gamma]$.
7. Let S be a subring of $\mathbb{R}$ which is a discrete set in the sense of Chapter 5 (4.3). Prove that $S = \mathbb{Z}$.
8. In each case, decide whether or not S is a subring of R.
   (a) S is the set of all rational numbers of the form a/b, where b is not divisible by 3, and $R = \mathbb{Q}$.
   (b) S is the set of functions which are linear combinations of the functions $\{1, \cos nt, \sin nt \mid n \in \mathbb{Z}\}$, and R is the set of all functions $\mathbb{R} \longrightarrow \mathbb{R}$.
   (c) (not commutative) S is the set of real matrices of the form $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$, and R is the set of all real 2×2 matrices.
9. In each case, decide whether the given structure forms a ring. If it is not a ring, determine which of the ring axioms hold and which fail:
   (a) U is an arbitrary set, and R is the set of subsets of U. Addition and multiplication of elements of R are defined by the rules $A + B = A \cup B$ and $A \cdot B = A \cap B$.
   (b) U is an arbitrary set, and R is the set of subsets of U. Addition and multiplication of elements of R are defined by the rules $A + B = (A \cup B) - (A \cap B)$ and $A \cdot B = A \cap B$.
   (c) R is the set of continuous functions $\mathbb{R} \longrightarrow \mathbb{R}$. Addition and multiplication are defined by the rules $[f + g](x) = f(x) + g(x)$ and $[f \circ g](x) = f(g(x))$.
10. Determine all rings which contain the zero ring as a subring.

11. Describe the group of units in each ring.
(a) $\mathbb{Z}/12\mathbb{Z}$ (b) $\mathbb{Z}/7\mathbb{Z}$ (c) $\mathbb{Z}/8\mathbb{Z}$ (d) $\mathbb{Z}/n\mathbb{Z}$
12. Prove that the units in the ring of Gauss integers are $\{\pm 1, \pm i\}$.
13. An element x of a ring R is called *nilpotent* if some power of x is zero. Prove that if x is nilpotent, then $1 + x$ is a unit in R.
14. Prove that the product set $R \times R'$ of two rings is a ring with component-wise addition and multiplication:

$$(a,a') + (b,b') = (a + b, a' + b') \quad \text{and} \quad (a,a')(b,b') = (ab, a'b').$$

This ring is called the *product ring*.

2. Formal Construction of Integers and Polynomials

1. Prove that every natural number n except 1 has the form m' for some natural number m.
2. Prove the following laws for the natural numbers.
(a) the commutative law for addition
(b) the associative law for multiplication
(c) the distributive law
(d) the cancellation law for addition: if $a + b = a + c$, then $b = c$
(e) the cancellation law for multiplication: if $ab = ac$, then $b = c$
3. The relation $<$ on $\mathbb{N}$ can be defined by the rule $a < b$ if $b = a + n$ for some n. Assume that the elementary properties of addition have been proved.
(a) Prove that if $a < b$, then $a + n < b + n$ for all n.
(b) Prove that the relation $<$ is transitive.
(c) Prove that if a, b are natural numbers, then precisely one of the following holds:

$$a < b, a = b, b < a.$$

(d) Prove that if $n \neq 1$, then $a < an$.
4. Prove the principle of *complete induction:* Let S be a subset of $\mathbb{N}$ with the following property: If n is a natural number such that $m \in S$ for every $m < n$, then $n \in S$. Then $S = \mathbb{N}$.
*5. Define the set $\mathbb{Z}$ of all integers, using two copies of $\mathbb{N}$ and an element representing zero, define addition and multiplication, and derive the fact that $\mathbb{Z}$ is a ring from the properties of addition and multiplication of natural numbers.
6. Let R be a ring. The set of all formal power series $p(t) = a_0 + a_1t + a_2t^2 + \cdots$, with $a_i \in R$, forms a ring which is usually denoted by $R[[t]]$. (By *formal power series* we mean that there is no requirement of convergence.)
(a) Prove that the formal power series form a ring.
(b) Prove that a power series $p(t)$ is invertible if and only if a_0 is a unit of R.
7. Prove that the units of the polynomial ring $\mathbb{R}[x]$ are the nonzero constant polynomials.

3. Homomorphisms and Ideals

1. Show that the inverse of a ring isomorphism $\varphi: R \longrightarrow R'$ is an isomorphism.
2. Prove or disprove: If an ideal I contains a unit, then it is the unit ideal.
3. For which integers n does $x^2 + x + 1$ divide $x^4 + 3x^3 + x^2 + 6x + 10$ in $\mathbb{Z}/n\mathbb{Z}[x]$?

4. Prove that in the ring $\mathbb{Z}[x]$, $(2) \cap (x) = (2x)$.
5. Prove the equivalence of the two definitions (3.11) and (3.12) of an ideal.
6. Is the set of polynomials $a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ such that 2^{k+1} divides a_k an ideal in $\mathbb{Z}[x]$?
7. Prove that every nonzero ideal in the ring of Gauss integers contains a nonzero integer.
8. Describe the kernel of the following maps.
   (a) $\mathbb{R}[x,y] \longrightarrow \mathbb{R}$ defined by $f(x,y) \rightsquigarrow f(0,0)$
   (b) $\mathbb{R}[x] \longrightarrow \mathbb{C}$ defomed by $f(x) \rightsquigarrow f(2+i)$
9. Describe the kernel of the map $\mathbb{Z}[x] \longrightarrow \mathbb{R}$ defined by $f(x) \rightsquigarrow f(1+\sqrt{2})$.
10. Describe the kernel of the homomorphism $\varphi: \mathbb{C}[x,y,z] \longrightarrow \mathbb{C}[t]$ defined by $\varphi(x) = t$, $\varphi(y) = t^2$, $\varphi(z) = t^3$.
11. (a) Prove that the kernel of the homomorphism $\varphi: \mathbb{C}[x,y] \longrightarrow \mathbb{C}[t]$ defined by $x \rightsquigarrow t^2$, $y \rightsquigarrow t^3$ is the principal ideal generated by the polynomial $y^2 - x^3$.
    (b) Determine the image of φ explicitly.
12. Prove the existence of the homomorphism (3.8).
13. State and prove an analogue of (3.8) when $\mathbb{R}$ is replaced by an arbitrary infinite field.
14. Prove that if two rings R, R′ are isomorphic, so are the polynomial rings $R[x]$ and $R'[x]$.
15. Let R be a ring, and let $f(y) \in R[y]$ be a polynomial in one variable with coefficients in R. Prove that the map $R[x,y] \longrightarrow R[x,y]$ defined by $x \rightsquigarrow x + f(y)$, $y \rightsquigarrow y$ is an automorphism of $R[x,y]$.
16. Prove that a polynomial $f(x) = \Sigma a_i x^i$ can be expanded in powers of $x - a$: $f(x) = \Sigma c_i (x-a)^i$, and that the coefficients c_i are polynomials in the coefficients a_i, with integer coefficients.
17. Let R, R' be rings, and let $R \times R'$ be their product. Which of the following maps are ring homomorphisms?
    (a) $R \longrightarrow R \times R'$, $r \rightsquigarrow (r, 0)$
    (b) $R \longrightarrow R \times R$, $r \rightsquigarrow (r, r)$
    (c) $R \times R' \longrightarrow R$, $(r_1, r_2) \rightsquigarrow r_1$
    (d) $R \times R \longrightarrow R$, $(r_1, r_2) \rightsquigarrow r_1 r_2$
    (e) $R \times R \longrightarrow R$, $(r_1, r_2) \rightsquigarrow r_1 + r_2$
18. (a) Is $\mathbb{Z}/(10)$ isomorphic to $\mathbb{Z}/(2) \times \mathbb{Z}/(5)$?
    (b) Is $\mathbb{Z}/(8)$ isomorphic to $\mathbb{Z}/(2) \times \mathbb{Z}/(4)$?
19. Let R be a ring of characteristic p. Prove that the map $R \longrightarrow R$ defined by $x \rightsquigarrow x^p$ is a ring homomorphism. This map is called the *Frobenius homomorphism*.
20. Determine all automorphisms of the ring $\mathbb{Z}[x]$.
21. Prove that the map $\mathbb{Z} \longrightarrow R$ (3.9) is compatible with multiplication of positive integers.
22. Prove that the characteristic of a field is either zero or a prime integer.
23. Let R be a ring of characteristic p. Prove that if a is nilpotent then $1 + a$ is *unipotent*, that is, some power of $1 + a$ is equal to 1.
24. (a) The *nilradical* N of a ring R is the set of its nilpotent elements. Prove that N is an ideal.
    (b) Determine the nilradicals of the rings $\mathbb{Z}/(12)$, $\mathbb{Z}/(n)$, and $\mathbb{Z}$.
25. (a) Prove Corollary (3.20).
    (b) Prove Corollary (3.22).

26. Determine all ideals of the ring $\mathbb{R}[[t]]$ of formal power series with real coefficients.

27. Find an ideal in the polynomial ring $F[x, y]$ in two variables which is not principal.

***28.** Let R be a ring, and let I be an ideal of the polynomial ring $R[x]$. Suppose that the lowest degree of a nonzero element of I is n and that I contains a monic polynomial of degree n. Prove that I is a principal ideal.

29. Let I, J be ideals of a ring R. Show by example that $I \cup J$ need not be an ideal, but show that $I + J = \{r \in R \mid r = x + y, \quad \text{with } x \in I, y \in J\}$ is an ideal. This ideal is called the *sum* of the ideals I, J.

30. (a) Let I, J be ideals of a ring R. Prove that $I \cap J$ is an ideal.
(b) Show by example that the set of products $\{xy \mid x \in I, y \in J\}$ need not be an ideal, but that the set of finite sums $\sum x_\nu y_\nu$ of products of elements of I and J is an ideal. This ideal is called the *product ideal.*
(c) Prove that $IJ \subset I \cap J$.
(d) Show by example that IJ and $I \cap J$ need not be equal.

31. Let I, J, J' be ideals in a ring R. Is it true that $I(J + J') = IJ + IJ'$?

***32.** If R is a noncommutative ring, the definition of an *ideal* is a set I which is closed under addition and such that if $r \in R$ and $x \in I$, then both rx and xr are in I. Show that the noncommutative ring of $n \times n$ real matrices has no proper ideal.

33. Prove or disprove: If $a^2 = a$ for all a in a ring R, then R has characteristic 2.

34. An element e of a ring S is called *idempotent* if $e^2 = e$. Note that in a product $R \times R'$ of rings, the element $e = (1, 0)$ is idempotent. The object of this exercise is to prove a converse.
(a) Prove that if e is idempotent, then $e' = 1 - e$ is also idempotent.
(b) Let e be an idempotent element of a ring S. Prove that the principal ideal eS is a ring, with identity element e. It will probably not be a subring of S because it will not contain 1 unless $e = 1$.
(c) Let e be idempotent, and let $e' = 1 - e$. Prove that S is isomorphic to the product ring $(eS) \times (e'S)$.

4. Quotient Rings and Relations in a Ring

1. Prove that the image of the homomorphism φ of Proposition (4.9) is the subring described in the proposition.

2. Determine the structure of the ring $\mathbb{Z}[x]/(x^2 + 3, p)$, where (a) $p = 3$, (b) $p = 5$.

3. Describe each of the following rings.
(a) $\mathbb{Z}[x]/(x^2 - 3, 2x + 4)$ (b) $\mathbb{Z}[i]/(2 + i)$

4. Prove Proposition (4.2).

5. Let R' be obtained from a ring R by introducing the relation $\alpha = 0$, and let $\psi: R \longrightarrow R'$ be the canonical map. Prove the following *universal property* for this construction: Let $\varphi: R \longrightarrow \tilde{R}$ be a ring homomorphism, and assume that $\varphi(\alpha) = 0$ in $\tilde{R}$. There is a unique homomorphism $\varphi': R' \longrightarrow \tilde{R}$ such that $\varphi' \circ \psi = \varphi$.

6. Let I, J be ideals in a ring R. Prove that the residue of any element of $I \cap J$ in R/IJ is nilpotent.

7. Let I, J be ideals of a ring R such that $I + J = R$.
(a) Prove that $IJ = I \cap J$.

*(b) Prove the *Chinese Remainder Theorem:* For any pair a, b of elements of R, there is an element x such that $x \equiv a$ (modulo I) and $x \equiv b$ (modulo J). [The notation $x \equiv a$ (modulo I) means $x - a \in I$.]

8. Let I, J be ideals of a ring R such that $I + J = R$ and $IJ = 0$.
(a) Prove that R is isomorphic to the product $(R/I) \times (R/J)$.
(b) Describe the idempotents corresponding to this product decomposition (see exercise 34, Section 3).

5. Adjunction of Elements

1. Describe the ring obtained from $\mathbb{Z}$ by adjoining an element α satisfying the two relations $2\alpha - 6 = 0$ and $\alpha - 10 = 0$.

2. Suppose we adjoin an element α to $\mathbb{R}$ satisfying the relation $\alpha^2 = 1$. Prove that the resulting ring is isomorphic to the product ring $\mathbb{R} \times \mathbb{R}$, and find the element of $\mathbb{R} \times \mathbb{R}$ which corresponds to α.

3. Describe the ring obtained from the product ring $\mathbb{R} \times \mathbb{R}$ by inverting the element $(2, 0)$.

4. Prove that the elements $1, t - \alpha, (t - \alpha)^2, \ldots, (t - \alpha)^{n-1}$ form a $\mathbb{C}$-basis for $\mathbb{C}[t]/((t - \alpha)^n)$.

5. Let α denote the residue of x in the ring $R' = \mathbb{Z}[x]/(x^4 + x^3 + x^2 + x + 1)$. Compute the expressions for $(\alpha^3 + \alpha^2 + \alpha)(\alpha + 1)$ and α^5 in terms of the basis $(1, \alpha, \alpha^2, \alpha^3, \alpha^4)$.

6. In each case, describe the ring obtained from $\mathbb{F}_2$ by adjoining an element α satisfying the given relation.
(a) $\alpha^2 + \alpha + 1 = 0$ (b) $\alpha^2 + 1 = 0$

7. Analyze the ring obtained from $\mathbb{Z}$ by adjoining an element α which satisfies the pair of relations $\alpha^3 + \alpha^2 + 1 = 0$ and $\alpha^2 + \alpha = 0$.

8. Let $a \in R$. If we adjoin an element α with the relation $\alpha = a$, we expect to get back a ring isomorphic to R. Prove that this is so.

9. Describe the ring obtained from $\mathbb{Z}/12\mathbb{Z}$ by adjoining an inverse of 2.

10. Determine the structure of the ring R' obtained from $\mathbb{Z}$ by adjoining element α satisfying each set of relations.
(a) $2\alpha = 6, 6\alpha = 15$ (b) $2\alpha = 6, 6\alpha = 18$ (c) $2\alpha = 6, 6\alpha = 8$

11. Let $R = \mathbb{Z}/(10)$. Determine the structure of the ring obtained by adjoining an element α satisfying each relation.
(a) $2\alpha - 6 = 0$ (b) $2\alpha - 5 = 0$

12. Let a be a unit in a ring R. Describe the ring $R' = R[x]/(ax - 1)$.

13. (a) Prove that the ring obtained by inverting x in the polynomial ring $R[x]$ is isomorphic to the ring of Laurent polynomials, as asserted in (5.9).
(b) Do the formal Laurent series $\sum_{-\infty}^{\infty} a_n x^n$ form a ring?

14. Let a be an element of a ring R, and let $R' = R[x]/(ax - 1)$ be the ring obtained by adjoining an inverse of a to R. Prove that the kernel of the map $R \longrightarrow R'$ is the set of elements $b \in R$ such that $a^n b = 0$ for some $n > 0$.

15. Let a be an element of a ring R, and let R' be the ring obtained from R by adjoining an inverse of a. Prove that R' is the zero ring if and only if a is nilpotent.

16. Let F be a field. Prove that the rings $F[x]/(x^2)$ and $F[x]/(x^2 - 1)$ are isomorphic if and only if F has characteristic 2.
17. Let $\overline{R} = \mathbb{Z}[x]/(2x)$. Prove that every element of $\overline{R}$ has a unique expression in the form $a_0 + a_1x + \cdots + a_nx^n$, where a_i are integers and $a_1,\ldots,a_n$ are either 0 or 1.

6. *Integral Domains and Fraction Fields*

1. Prove that a subring of an integral domain is an integral domain.
2. Prove that an integral domain with finitely many elements is a field.
3. Let R be an integral domain. Prove that the polynomial ring $R[x]$ is an integral domain.
4. Let R be an integral domain. Prove that the invertible elements of the polynomial ring $R[x]$ are the units in R.
5. Is there an integral domain containing exactly 10 elements?
6. Prove that the field of fractions of the formal power series ring $F[[x]]$ over a field F is obtained by inverting the single element x, and describe the elements of this field as certain power series with negative exponents.
7. Carry out the verification that the equivalence classes of fractions from an integral domain form a field.
8. A semigroup S is a set with an associative law of composition having an identity element. Let S be a commutative semigroup which satisfies the cancellation law: $ab = ac$ implies $b = c$. Use fractions to prove that S can be embedded into a group.

*9. A subset S of an integral domain R which is closed under multiplication and which does not contain 0 is called a *multiplicative set*. Given a multiplicative set S, we define S-fractions to be elements of the form a/b, where $b \in S$. Show that the equivalence classes of S-fractions form a ring.

7. *Maximal Ideals*

1. Prove that the maximal ideals of the ring of integers are the principal ideals generated by prime integers.
2. Determine the maximal ideals of each of the following.
   (a) $\mathbb{R} \times \mathbb{R}$ (b) $\mathbb{R}[x]/(x^2)$ (c) $\mathbb{R}[x]/(x^2 - 3x + 2)$ (d) $\mathbb{R}[x]/(x^2 + x + 1)$
3. Prove that the ideal $(x + y^2, y + x^2 + 2xy^2 + y^4)$ in $\mathbb{C}[x,y]$ is a maximal ideal.
4. Let R be a ring, and let I be an ideal of R. Let M be an ideal of R containing I, and let $\overline{M} = M/I$ be the corresponding ideal of $\overline{R}$. Prove that M is maximal if and only if $\overline{M}$ is.
5. Let I be the principal ideal of $\mathbb{C}[x,y]$ generated by the polynomial $y^2 + x^3 - 17$. Which of the following sets generate maximal ideals in the quotient ring $R = \mathbb{C}[x,y]/I$?
   (a) $(x - 1, y - 4)$ (b) $(x + 1, y + 4)$ (c) $(x^3 - 17, y^2)$
6. Prove that the ring $\mathbb{F}_5[x]/(x^2 + x + 1)$ is a field.
7. Prove that the ring $\mathbb{F}_2[x]/(x^3 + x + 1)$ is a field, but that $\mathbb{F}_3[x]/(x^3 + x + 1)$ is not a field.
8. Let $R = \mathbb{C}[x_1,\ldots,x_n]/I$ be a quotient of a polynomial ring over $\mathbb{C}$, and let M be a maximal ideal of R. Prove that $R/M \approx \mathbb{C}$.
9. Define a bijective correspondence between maximal ideals of $\mathbb{R}[x]$ and points in the upper half plane.

10. Let R be a ring, with M an ideal of R. Suppose that every element of R which is not in M is a unit of R. Prove that M is a maximal ideal and that moreover it is the only maximal ideal of R.

11. Let P be an ideal of a ring R. Prove that $\overline{R} = R/P$ is an integral domain if and only if $P \neq R$, and that if $a, b \in R$ and $ab \in P$, then $a \in P$ or $b \in P$. (An ideal P satisfying these conditions is called a *prime ideal.*)

12. Let $\varphi: R \longrightarrow R'$ be a ring homomorphism, and let P' be a prime ideal of R'.
(a) Prove that $\varphi^{-1}(P')$ is a prime ideal of R.
(b) Give an example in which P' is a maximal ideal, but $\varphi^{-1}(P')$ is not maximal.

***13.** Let R be an integral domain with fraction field F, and let P be a prime ideal of R. Let R_P be the subset of F defined by

$$R_P = \{a/d \mid a, d \in R, d \notin P\}.$$

This subset is called the *localization of R at P*.
(a) Prove that R_P is a subring of F.
(b) Determine all maximal ideals of R_P.

14. Find an example of a "ring without unit element" and an ideal not contained in a maximal ideal.

8. Algebraic Geometry

1. Determine the points of intersection of the two complex plane curves in each of the following.
(a) $y^2 - x^3 + x^2 = 1, \quad x + y = 1$
(b) $x^2 + xy + y^2 = 1, \quad x^2 + 2y^2 = 1$
(c) $y^2 = x^3, \quad xy = 1$
(d) $x + y + y^2 = 0, \quad x - y + y^2 = 0$
(e) $x + y^2 = 0, \quad y + x^2 + 2xy^2 + y^4 = 0$

2. Prove that two quadratic polynomials f, g in two variables have at most four common zeros, unless they have a nonconstant factor in common.

3. Derive the Hilbert Nullstellensatz from its classical form (8.7).

4. Let U, V be varieties in $\mathbb{C}^n$. Prove that $U \cup V$ and $U \cap V$ are varieties.

5. Let $f_1, \ldots, f_r;\ g_1, \ldots, g_s \in \mathbb{C}[x_1, \ldots, x_n]$, and let U, V be the zeros of $\{f_1, \ldots, f_r\}$, $\{g_1, \ldots, g_s\}$ respectively. Prove that if U and V do not meet, then $(f_1, \ldots, f_r; g_1, \ldots, g_s)$ is the unit ideal.

6. Let $f = f_1 \cdots f_m$ and $g = g_1 \cdots g_n$, where f_i, g_j are irreducible polynomials in $\mathbb{C}[x, y]$. Let $S_i = \{f_i = 0\}$ and $T_j = \{g_j = 0\}$ be the Riemann surfaces defined by these polynomials, and let V be the variety $f = g = 0$. Describe V in terms of S_i, T_j.

7. Prove that the variety defined by a set $\{f_1, \ldots, f_r\}$ of polynomials depends only on the ideal $(f_1, \ldots, f_r)$ they generate.

8. Let R be a ring containing $\mathbb{C}$ as subring.
(a) Show how to make R into a vector space over $\mathbb{C}$.
(b) Assume that R is a finite-dimensional vector space over $\mathbb{C}$ and that R contains exactly one maximal ideal M. Prove that M is the *nilradical* of R, that is, that M consists precisely of its nilpotent elements.

9. Prove that the complex conic $xy = 1$ is homeomorphic to the plane, with one point deleted.

10. Prove that every variety in $\mathbb{C}^2$ is the union of finitely many points and algebraic curves.

11. The three polynomials $f_1 = x^2 + y^2 - 1, f_2 = x^2 - y + 1$, and $f_3 = xy - 1$ generate the unit ideal in $\mathbb{C}[x,y]$. Prove this in two ways: (i) by showing that they have no common zeros, and (ii) by writing 1 as a linear combination of f_1, f_2, f_3, with polynomial coefficients.

12. **(a)** Determine the points of intersection of the algebraic curve S: $y^2 = x^3 - x^2$ and the line L: $y = \lambda x$.
(b) Parametrize the points of S as a function of λ.
(c) Relate S to the complex λ-plane, using this parametrization.

***13.** The *radical* of an ideal I is the set of elements $r \in R$ such that some power of r is in I.
(a) Prove that the radical of I is an ideal.
(b) Prove that the varieties defined by two sets of polynomials $\{f_1,\dots,f_r\},\{g_1,\dots,g_s\}$ are equal if and only if the two ideals $(f_1,\dots,f_r),(g_1,\dots,g_s)$ have the same radicals.

***14.** Let $R = \mathbb{C}[x_1,\dots,x_n]/(f_1,\dots,f_m)$. Let A be a ring containing $\mathbb{C}$ as subring. Find a bijective correspondence between the following sets:
(i) homomorphisms $\varphi: R \longrightarrow A$ which restrict to the identity on $\mathbb{C}$, and
(ii) n-tuples $a = (a_1,\dots,a_n)$ of elements of A which solve the system of equations $f_1 = \dots = f_m = 0$, that is, such that $f_i(a) = 0$ for $i = 1,\dots,m$.

Miscellaneous Exercises

1. Let F be a field, and let K denote the vector space F^2. Define multiplication by the rules $(a_1,a_2)\cdot(b_1,b_2) = (a_1b_1 - a_2b_2, a_1b_2 + a_2b_1)$.
(a) Prove that this law and vector addition make K into a ring.
(b) Prove that K is a field if and only if there is no element in F whose square is -1.
(c) Assume that -1 is a square in F and that F does not have characteristic 2. Prove that K is isomorphic to the product ring $F \times F$.

2. **(a)** We can define the derivative of an arbitrary polynomial $f(x)$ with coefficients in a ring R by the calculus formula $(a_nx^n + \cdots + a_1x + a_0)' = na_nx^{n-1} + \cdots + 1a_1$. The integer coefficients are interpreted in R using the homomorphism (3.9). Prove the product formula $(fg)' = f'g + fg'$ and the chain rule $(f \circ g)' = (f' \circ g)g'$.
(b) Let $f(x)$ be a polynomial with coefficients in a field F, and let α be an element of F. Prove that α is a multiple root of f if and only if it is a common root of f and of its derivative f'.
(c) Let $F = \mathbb{F}_5$. Determine whether or not the following polynomials have multiple roots in F: $x^{15} - x$, $x^{15} - 2x^5 + 1$.

3. Let R be a set with two laws of composition satisfying all the ring axioms except the commutative law for addition. Prove that this law holds by expanding the product $(a + b)(c + d)$ in two ways using the distributive law.

4. Let R be a ring. Determine the units in the polynomial ring $R[x]$.

5. Let R denote the set of sequences $a = (a_1,a_2,a_3,\dots)$ of real numbers which are eventually constant: $a_n = a_{n+1} = \dots$ for sufficiently large n. Addition and multiplication are component-wise; that is, addition is vector addition and $ab = (a_1b_1, a_2b_2,\dots)$.
(a) Prove that R is a ring.
(b) Determine the maximal ideals of R.

6. **(a)** Classify rings R which contain $\mathbb{C}$ and have dimension 2 as vector space over $\mathbb{C}$.
***(b)** Do the same as (a) for dimension 3.

*7. Consider the map φ: $\mathbb{C}[x,y] \longrightarrow \mathbb{C}[x] \times \mathbb{C}[y] \times \mathbb{C}[t]$ defined by $f(x,y) \rightsquigarrow (f(x,0), f(0,y), f(t,t))$. Determine the image of φ explicitly.

8. Let S be a subring of a ring R. The *conductor* C of S in R is the set of elements $\alpha \in R$ such that $\alpha R \subset S$.
   (a) Prove that C is an ideal of R and also an ideal of S.
   (b) Prove that C is the largest ideal of S which is also an ideal of R.
   (c) Determine the conductor in each of the following three cases:
   (i) $R = \mathbb{C}[t]$, $S = \mathbb{C}[t^2, t^3]$;
   (ii) $R = \mathbb{Z}[\zeta]$, $\zeta = \frac{1}{2}(-1 + \sqrt{-3})$, $S = \mathbb{Z}[\sqrt{-3}]$;
   (iii) $R = \mathbb{C}[t, t^{-1}]$, $S = \mathbb{C}[t]$.

9. A *line* in $\mathbb{C}^2$ is the locus of a linear equation L: $\{ax + by + c = 0\}$. Prove that there is a unique line through two points $(x_0, y_0), (x_1, y_1)$, and also that there is a unique line through a point (x_0, y_0) with a given tangent direction (u_0, v_0).

10. An algebraic curve C in $\mathbb{C}^2$ is called *irreducible* if it is the locus of zeros of an irreducible polynomial $f(x,y)$—one which can not be factored as a product of nonconstant polynomials. A point $p \in C$ is called a *singular point* of the curve if $\partial f/\partial x = \partial f/\partial y = 0$ at p. Otherwise p is a *nonsingular* point. Prove that an irreducible curve has only finitely many singular points.

11. Let L: $ax + by + c = 0$ be a line and C: $\{f = 0\}$ a curve in $\mathbb{C}^2$. Assume that $b \neq 0$. Then we can use the equation of the line to eliminate y from the equation $f(x,y) = 0$ of C, obtaining a polynomial $g(x)$ in x. Show that its roots are the x-coordinates of the intersection points.

12. With the notation as in the preceding problem, the *multiplicity of intersection* of L and C at a point $p = (x_0, y_0)$ is the multiplicity of x_0 as a root of $g(x)$. The line is called a *tangent line* to C at p if the multiplicity of intersection is at least 2. Show that if p is a nonsingular point of C, then there is a unique tangent line at (x_0, y_0), and compute it.

13. Show that if p is a singular point of a curve C, then the multiplicity of intersection of every line through p is at least 2.

14. The *degree* of an irreducible curve C: $\{f = 0\}$ is defined to be the degree of the irreducible polynomial f.
   (a) Prove that a line L meets C in at most d points, unless $C = L$.
   *(b) Prove that there exist lines which meet C in precisely d points.

15. Determine the singular points of $x^3 + y^3 - 3xy = 0$.

*16. Prove that an irreducible cubic curve can have at most one singular point.

*17. A nonsingular point p of a curve C is called a *flex point* if the tangent line L to C at p has an intersection of multiplicity at least 3 with C at p.
   (a) Prove that the flex points are the nonsingular points of C at which the *Hessian*

$$\det \begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} & \dfrac{\partial f}{\partial x} \\ \dfrac{\partial^2 f}{\partial x \partial y} & \dfrac{\partial^2 f}{\partial y^2} & \dfrac{\partial f}{\partial y} \\ \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} & f \end{bmatrix}$$

   vanishes.
   (b) Determine the flex points of the cubic curves $y^2 - x^3$ and $y^2 - x^3 + x^2$.

*18. Let C be an irreducible cubic curve, and let L be a line joining two flex points of C. Prove that if L meets C in a third point, then that point is also a flex.

19. Let $U = \{f_i(x_1,\ldots,x_m) = 0\}$, $V = \{g_j(y_1,\ldots,y_n) = 0\}$ be two varieties. Show that the variety defined by the equations $\{f_i(x) = 0, g_j(y) = 0\}$ in $\mathbb{C}^{m+n}$ is the product set $U \times V$.

20. Prove that the locus $y = \sin x$ in $\mathbb{R}^2$ doesn't lie on any algebraic curve.

*21. Let f, g be polynomials in $\mathbb{C}[x,y]$ with no common factor. Prove that the ring $R = \mathbb{C}[x,y]/(f, g)$ is a finite-dimensional vector space over $\mathbb{C}$.

22. **(a)** Let s, c denote the functions $\sin x$, $\cos x$ on the real line. Prove that the ring $\mathbb{R}[s,c]$ they generate is an integral domain.
   (b) Let $K = \mathbb{R}(s,c)$ denote the field of fractions of $\mathbb{R}[s,c]$. Prove that the field K is isomorphic to the field of rational functions $\mathbb{R}(x)$.

*23. Let $f(x)$, $g(x)$ be polynomials with coefficients in a ring R with $f \neq 0$. Prove that if the product $f(x)g(x)$ is zero, then there is a nonzero element $c \in R$ such that $cg(x) = 0$.

*24. Let X denote the closed unit interval $[0, 1]$, and let R be the ring of continuous functions $X \longrightarrow \mathbb{R}$.
   (a) Prove that a function f which does not vanish at any point of X is invertible in R.
   (b) Let $f_1,\ldots,f_n$ be functions with no common zero on X. Prove that the ideal generated by these functions is the unit ideal. (Hint: Consider $f_1^2 + \cdots + f_n^2$.)
   (c) Establish a bijective correspondence between maximal ideals of R and points on the interval.
   (d) Prove that the maximal ideals containing a function f correspond to points of the interval at which $f = 0$.
   (e) Generalize these results to functions on an arbitrary compact set X in $\mathbb{R}^k$.
   (f) Describe the situation in the case $X = \mathbb{R}$.

Chapter 11

Factorization

Rien n'est beau que le vrai.

Hermann Minkowski

1. FACTORIZATION OF INTEGERS AND POLYNOMIALS

This chapter is a study of division in rings. Because it is modeled on properties of the ring of integers, we will begin by reviewing these properties. Some have been used without comment in earlier chapters of the book, and some have already been proved.

The property from which all others follow is division with remainder: If a, b are integers and $a \neq 0$, there exist integers q, r so that

$$b = aq + r, \tag{1.1}$$

and $0 \leq r < |a|$. This property is often stated only for positive integers, but we allow a and b to take on negative values too. That is why we use the absolute value $|a|$ to bound the remainder. The proof of the existence of (1.1) is a simple induction argument.

We've already seen some of the most important consequences of division with remainder, but let us recall them. In Chapter 10, we saw that every subgroup of $\mathbb{Z}^+$ is an ideal and that every ideal of $\mathbb{Z}$ is principal, that is, it has the form $d\mathbb{Z}$ for some integer $d \geq 0$. As was proved in Chapter 2 (2.6), this implies that a greatest common divisor of a pair of integers a, b exists and that it is an integer linear combination of a and b. If a and b have no factor in common other than ± 1, then 1 is a linear combination of a and b with integer coefficients:

$$ra + sb = 1, \tag{1.2}$$

for some $r, s \in \mathbb{Z}$. This implies the fundamental property of prime integers, which was proved in Chapter 3 (2.8). We restate it here:

(1.3) **Proposition.** Let p be a prime integer, and let a, b be integers. If p divides the product ab, then p divides a or b. □

(1.4) **Theorem.** *Fundamental Theorem of Arithmetic:* Every integer $a \neq 0$ can be written as a product

$$a = cp_1 \cdots p_k,$$

where $c = \pm 1$, the p_i are positive prime integers, and $k \geq 0$. This expression is unique except for the ordering of the prime factors.

Proof. First, a prime factorization exists. To prove this, it is enough to consider the case that a is greater than 1. By induction on a, we may assume the existence proved for all positive integers $b < a$. Either a is prime, in which case the product has one factor, or there is a proper divisor $b \neq a$. Then $a = bb'$ and also $b' \neq a$. Both b and b' are smaller than a, and by induction they can be factored into primes. Setting their factorizations side by side gives a factorization of a.

Second, the factorization is unique. Suppose that

$$\pm p_1 \cdots p_n = a = \pm q_1 \cdots q_m.$$

The signs certainly agree. We apply (1.3), with $p = p_1$. Since p_1 divides the product $q_1 \cdots q_m$, it divides some q_i, say q_1. Since q_1 is prime, $p_1 = q_1$. Cancel p_1 and proceed by induction. □

The structure of the ring of integers is closely analogous to that of a polynomial ring $F[x]$ in one variable over a field. Whenever a property of one of these rings is derived, we should try to find an analogous property of the other. We have already discussed division with remainder for polynomials in Chapter 10, and we have seen that every ideal of the polynomial ring $F[x]$ is principal [Chapter 10 (3.21)].

A polynomial $p(x)$ with coefficients in a field F is called *irreducible* if it is not constant and if its only divisors of lower degree in $F[x]$ are constants. This means that the only way that p can be written as a product of two polynomials is $p = cp_1$, where c is a constant and p_1 is a constant multiple of p. The irreducible polynomials are analogous to prime integers. It is customary to normalize them by factoring out their leading coefficients, so that they become monic.

The proof of the following theorem is similar to the proof of the analogous statements for the ring of integers:

(1.5) **Theorem.** Let F be a field, and let $F[x]$ denote the polynomial ring in one variable over F.

(a) If two polynomials f, g have no common nonconstant factor, then there are polynomials $r, s \in F[x]$ such that $rf + sg = 1$.

(b) If an irreducible polynomial $p \in F[x]$ divides a product fg, then p divides one of the factors f or g.

(c) Every nonzero polynomial $f \in F[x]$ can be written as a product

$$f = cp_1 \cdots p_k,$$

where c is a nonzero constant, the p_i are monic irreducible polynomials in $F[x]$, and $k \geq 0$. This factorization is unique, except for the ordering of the terms. □

The constant factor c which appears in the third part of this theorem is analogous to the factor ± 1 in (1.4). These are the units in their respective rings. The unit factors are there because we normalized primes to be positive, and irreducible polynomials to be monic. We can allow negative primes or nonmonic irreducible polynomials if we wish. The unit factor can then be absorbed, if $k > 0$. But this complicates the statement of uniqueness slightly.

(1.6) **Examples.** Over the complex numbers, every polynomial of positive degree has a root α and therefore has a divisor of the form $x - \alpha$. So the irreducible polynomials are linear, and the irreducible factorization of a polynomial has the form

$$f(x) = c(x - \alpha_1) \cdots (x - \alpha_n), \tag{1.7}$$

where α_i are the roots of $f(x)$, repeated as necessary. The uniqueness of this factorization is not surprising.

When $F = \mathbb{R}$, there are two classes of irreducible polynomials: linear polynomials and irreducible quadratic polynomials. A real quadratic polynomial $x^2 + bx + c$ is irreducible if and only if its discriminant $b^2 - 4c$ is negative, in which case it has a pair of complex conjugate roots. The fact that every irreducible polynomial over the complex numbers is linear implies that no higher-degree polynomial is irreducible over the reals. Suppose that a polynomial $f(x)$ has real coefficients a_i and that α is a complex, nonreal root of $f(x)$. Then the complex conjugate $\overline{\alpha}$ is different from α and is also a root. For, since f is a real polynomial, its coefficients a_i satisfy the relation $a_i = \overline{a}_i$. Then

$$f(\overline{\alpha}) = a_n\overline{\alpha}^n + \cdots + a_1\overline{\alpha} + a_0 = \overline{a}_n\overline{\alpha}^n + \cdots + \overline{a}_1\overline{\alpha} + \overline{a}_0 = \overline{f(\alpha)} = \overline{0} = 0.$$

The quadratic polynomial $g(x) = (x - \alpha)(x - \overline{\alpha}) = x^2 - (\alpha + \overline{\alpha})x + \alpha\overline{\alpha}$ has real coefficients $-(\alpha + \overline{\alpha})$ and $\alpha\overline{\alpha}$, and both of its linear factors appear on the right side of the complex factorization (1.7) of $f(x)$. Thus $g(x)$ divides $f(x)$. So the factorization of $f(x)$ into irreducible real polynomials is obtained by grouping conjugate pairs in the complex factorization. □

Factorization of polynomials is more complicated for polynomials with rational coefficients than for real or complex polynomials, because there exist irreducible polynomials in $\mathbb{Q}[x]$ of arbitrary degree. For example, $x^5 - 3x^4 + 3$ is irreducible in $\mathbb{Q}[x]$. We will see more examples in Section 4. Neither the form of the irreducible factorization nor its uniqueness is intuitively clear for rational polynomials.

For future reference, we note the following elementary fact:

(1.8) **Proposition.** Let F be a field, and let $f(x)$ be a polynomial of degree n with coefficients in F. Then f has at most n roots in F.

Proof. An element $\alpha \in F$ is a root of f if and only if $x - \alpha$ divides f [Chapter 10 (3.20)]. If so, then we can write $f(x) = (x - \alpha)q(x)$, where $q(x)$ is a polynomial of degree $n - 1$. If β is another root of f, then $f(\beta) = (\beta - \alpha)q(\beta) = 0$. Since F is a field, the product of nonzero elements of F is not zero. So one of the two elements $\beta - \alpha$, $q(\beta)$ is zero. In the first case $\beta = \alpha$, and in the second case β is one of the roots of $q(x)$. By induction on n, we may assume that $q(x)$ has at most $n - 1$ roots in F. Then there are at most n possibilities for β. □

The fact that F is a field is crucial to Theorem (1.5) and to Proposition (1.8), as the following example shows. Let R be the ring $\mathbb{Z}/8\mathbb{Z}$. Then in the polynomial ring $R[x]$, we have

$$x^2 - 1 = (x + 1)(x - 1) = (x + 3)(x - 3).$$

The polynomial $x^2 - 1$ has four roots modulo 8, and its factorization into irreducible polynomials is not unique.

2. UNIQUE FACTORIZATION DOMAINS, PRINCIPAL IDEAL DOMAINS, AND EUCLIDEAN DOMAINS

Having seen that factorization of polynomials is analogous to factorization of integers, it is natural to ask whether other rings can have such properties. Relatively few such rings exist, but the ring of Gauss integers is one interesting example. This section explores ways in which various parts of the theory can be extended.

We begin by introducing the terminology used in studying factorization. It is natural to assume that the given ring R is an integral domain, so that the Cancellation Law is available, and we will make this assumption throughout. We say that an element a *divides* another element b (abbreviated $a \mid b$) if $b = aq$ for some $q \in R$. The element a is a *proper divisor* of b if $b = aq$ for some $q \in R$ and if neither a nor q is a unit. A nonzero element a of R is called *irreducible* if it is not a unit and if it has no proper divisor. Two elements a, a' are called *associates* if each divides the other. It is easily seen that a, a' are associates if and only if they differ by a unit factor, that is, if $a' = ua$ for some unit u.

The concepts of divisor, unit, and associate can be interpreted in terms of the principal ideals generated by the elements. Remember that an ideal I is called *principal* if it is generated by a single element:

$$I = (a). \tag{2.1}$$

Keep in mind the fact that (a) consists of all elements which are multiples of a, that is, which are divisible by a. Then

$$\text{(2.2)} \qquad \begin{aligned} u \text{ is a unit} &\Leftrightarrow (u) = (1) \\ a \text{ and } a' \text{ are associates} &\Leftrightarrow (a) = (a') \\ a \text{ divides } b &\Leftrightarrow (a) \supset (b) \\ a \text{ is a proper divisor of } b &\Leftrightarrow (1) > (a) > (b). \end{aligned}$$

The proof of these equivalences is straightforward, and we omit it.

Now suppose that we hope for a theorem analogous to the Fundamental Theorem of Arithmetic in an integral domain R. We may divide the statement of the theorem into two parts. First, a given element a must be a product of irreducible elements, and second, this product must be essentially unique.

Consider the first part. We assume that our element a is not zero and not a unit; otherwise we have no hope of writing it as a product of irreducible elements. Then we attempt to factor a, proceeding as follows: If a is irreducible itself, we are done. If not, then a has a proper factor, so it decomposes in some way as a product, $a = a_1 b_1$, where neither a_1 nor b_1 is a unit. We continue factoring a_1 and b_1 if possible, and we hope that this procedure terminates; in other words, we hope that after a finite number of steps all the factors are irreducible. The condition that this procedure always terminates has a neat description in terms of principal ideals:

(2.3) **Proposition.** Let R be an integral domain. The following conditions are equivalent:

(a) For every nonzero element a of R which is not a unit, the process of factoring a terminates after finitely many steps and results in a factorization $a = b_1 \cdots b_k$ of a into irreducible elements of R.

(b) R does not contain an infinite increasing chain of principal ideals

$$(a_1) < (a_2) < (a_3) < \dots .$$

Proof. Suppose that R contains an infinite increasing sequence $(a_1) < (a_2) < \dots$. Then $(a_n) < (1)$ for every n, because $(a_n) < (a_{n+1}) \subset (1)$. Since $(a_{n-1}) < (a_n)$, a_n is a proper divisor of a_{n-1}, say $a_{n-1} = a_n b_n$ where a_n, b_n are not units. This provides a nonterminating sequence of factorizations of a_1: $a_1 = a_2 b_2 = a_3 b_3 b_2 = a_4 b_4 b_3 b_2 \dots$. Conversely, such a sequence of factorizations gives us an increasing chain of ideals. □

The second condition of this proposition is often called the *ascending chain condition* for principal ideals. However, to emphasize the factorization property, we will say that *existence of factorizations* holds in R if the equivalent conditions of the proposition are true.

It is easy to describe domains in which existence of factorizations fails. One example is obtained by adjoining all 2^k-th roots of x_1 to the polynomial ring $F[x_1]$:

$$\text{(2.4)} \qquad R = F[x_1, x_2, x_3, \dots],$$

with the relations $x_2^2 = x_1, x_3^2 = x_2, x_4^2 = x_3$, and so on. We can factor the element x_1 indefinitely in this ring, and correspondingly there is an infinite chain $(x_1) < (x_2) < \ldots$ of principal ideals.

It turns out that we need infinitely many generators for a ring to make an example such as the one just given, so we will rarely encounter such rings. In practice, the second part of the Fundamental Theorem is the one which gives the most trouble. Factorization into irreducible elements will usually be possible, but it will not be unique.

Units in a ring complicate the statement of uniqueness. It is clear that unit factors should be disregarded, since there is no end to the possibility of adding unit factors in pairs uu^{-1}. For the same reason, *associate* factors should be considered equivalent. The units in the ring of integers are ± 1, and in this ring it was natural to normalize irreducible elements (primes) to be positive; similarly, we may normalize irreducible polynomials by insisting that they be monic. We don't have a reasonable way to normalize elements of an arbitrary integral domain, so we will allow some ambiguity. It is actually neater to work with *principal ideals* than with elements: Associates generate the same principal ideal. However, it isn't too cumbersome to use elements here, and we will stay with them. The importance of ideals will become clear in the later sections of this chapter.

We will call an integral domain R a *unique factorization domain* if it has the following properties:

(2.5)

(i) Existence of factorizations is true for R. In other words, the process of factoring a nonzero element a which is not a unit terminates after finitely many steps and yields a factorization $a = p_1 \cdots p_m$, where each p_i is irreducible.

(ii) The irreducible factorization of an element is unique in the following sense: If a is factored in two ways into irreducible elements, say $a = p_1 \cdots p_m = q_1 \cdots q_n$, then $m = n$, and with suitable ordering of the factors, p_i is an associate of q_i for each i.

So in the statement of uniqueness, associate factorizations are considered equivalent.

Here is an example in which uniqueness of factorization is not true. The ring is the integral domain

$$R = \mathbb{Z}[\sqrt{-5}]. \tag{2.6}$$

It consists of all complex numbers of the form $a + b\sqrt{-5}$, where $a, b \in \mathbb{Z}$. The units in this ring are ± 1, and the integer 6 has two essentially different factorizations in R:

$$6 = 2 \cdot 3 = (1+\sqrt{-5})(1-\sqrt{-5}). \tag{2.7}$$

It is not hard to show that all four terms $2, 3, 1+\sqrt{-5}, 1-\sqrt{-5}$ are irreducible elements of R. Since the units are ± 1, the associates of 2 are 2 and -2. So 2 is not an associate of $1 \pm \sqrt{-5}$, which shows that the two factorizations are essentially different and hence that R is not a unique factorization domain.

The crucial property of prime integers is that if a prime divides a product, it divides one of the factors. We will call an element p of an integral domain R *prime* if it has these properties: p is not zero and not a unit, and if p divides a product of elements of R, it divides one of the factors. These are the properties from which uniqueness of the factorization is derived.

(2.8) **Proposition.** Let R be an integral domain. Suppose that existence of factorizations holds in R. Then R is a unique factorization domain if and only if every irreducible element is prime.

The proof is a simple extension of the arguments used in (1.3) and (1.4); we leave it as an exercise. □

It is important to distinguish between the two concepts of irreducible element and prime element. They are equivalent in unique factorization domains, but most rings contain irreducible elements which are not prime. For instance, in the ring $R = \mathbb{Z}[\sqrt{-5}]$ considered above, the element 2 has no proper factor, so it is irreducible. It is not prime because, though it divides the product $6 = (1+\sqrt{-5})(1-\sqrt{-5})$, it does not divide either factor.

Since irreducible elements in a unique factorization domain are prime, the phrases *irreducible factorization* and *prime factorization* are synonymous. We can use them interchangeably when we are working in a unique factorization domain, but not otherwise.

There is a simple way of deciding whether an element a divides another element b in a unique factorization domain, in terms of their irreducible (or prime) factorizations.

(2.9) **Proposition.** Let R be a unique factorization domain, and let $a = p_1 \cdots p_r$, $b = q_1 \cdots q_s$ be given prime factorizations of two elements of R. Then a divides b in R if and only if $s \geq r$, and with a suitable ordering of the factors q_i of b, p_i is an associate of q_i for $i = 1,\ldots,r$. □

(2.10) **Corollary.** Let R be a unique factorization domain, and let a, b be elements of R which are not both zero. There exists a *greatest common divisor* d of a, b, with the following properties:

(i) d divides a and b;

(ii) if an element e of R divides a and b, then e divides d. □

It follows immediately from the second condition that any two greatest common divisors of a, b are associates. However, *the greatest common divisor need not have the form $ra + sb$*. For example, we will show in the next section that the integer polynomial ring $\mathbb{Z}[x]$ is a unique factorization domain [see (3.8)]. In this ring, the elements 2 and x have greatest common divisor 1, but 1 is not a linear combination of these elements with integer polynomial coefficients.

Another important property of the ring of integers is that every ideal of $\mathbb{Z}$ is principal. An integral domain in which every ideal is principal is called a *principal ideal domain*.

(2.11) **Proposition.**

(a) In an integral domain, a prime element is irreducible.
(b) In a principal ideal domain, an irreducible element is prime.

We leave the proofs of (2.9–2.11) as exercises. □

(2.12) **Theorem.** A principal ideal domain is a unique factorization domain.

Proof. Suppose that R is a principal ideal domain. Then every irreducible element of R is prime. So according to Proposition (2.8), we need only prove the existence of factorizations for R. By Proposition (2.3), this is equivalent to showing that R contains no infinite increasing chain of principal ideals. We argue by contradiction. Suppose that $(a_1) < (a_2) < (a_3) < \dots$ is such a chain.

(2.13) **Lemma.** Let R be any ring. The union of an increasing chain of ideals $I_1 \subset I_2 \subset I_3 \subset \dots$ is an ideal.

Proof. Let I denote the union of the chain. If u, v are in I, then they are in I_n for some n. Then $u + v$ and ru are also in I_n; hence they are in I. □

We apply this lemma to the union I of our chain of principal ideals and use the hypothesis that R is a principal ideal domain to conclude that I is principal, say $I = (b)$. Now since b is in the union of the ideals (a_n), it is in one of these ideals. But if $b \in (a_n)$, then $(b) \subset (a_n)$, and on the other hand $(a_n) \subset (a_{n+1}) \subset (b)$. Therefore $(a_n) = (a_{n+1}) = (b)$. This contradicts the assumption that $(a_n) < (a_{n+1})$, and this contradiction completes the proof. □

The converse of Theorem (2.12) is not true. The ring $\mathbb{Z}[x]$ of integer polynomials is a unique factorization domain [see (3.8)], but it is not a principal ideal domain.

(2.14) **Proposition.**

(a) Let p be a nonzero element of a principal ideal domain R. Then $R/(p)$ is a field if and only if p is irreducible.
(b) The maximal ideals are the principal ideals generated by irreducible elements.

Proof. Since an ideal M is maximal if and only if R/M is a field, the two parts are equivalent. We will prove the second part. A principal ideal (a) contains another principal ideal (b) if and only if a divides b. The only divisors of an irreducible element p are the units and the associates of p. Therefore the only principal ideals which contain (p) are (p) and (1). Since every ideal of R is principal, this shows that an irreducible element generates a maximal ideal. Conversely, let b be a polynomial

having a proper factorization $b = aq$, where neither a nor q is a unit. Then $(b) < (a) < (1)$, and this shows that (b) is not maximal. □

Let us now abstract the procedure of division with remainder. To do so, we need a notion of *size* of an element of a ring. Appropriate measures are

(2.15)
$$\textit{absolute value}, \text{ if } R = \mathbb{Z},$$
$$\textit{degree of a polynomial}, \text{ if } R = F[x],$$
$$(\textit{absolute value})^2, \text{ if } R = \mathbb{Z}[i].$$

In general, a *size function* on an integral domain R will be any function

(2.16)
$$\sigma: R-\{0\} \longrightarrow \{0, 1, 2, \ldots\}$$

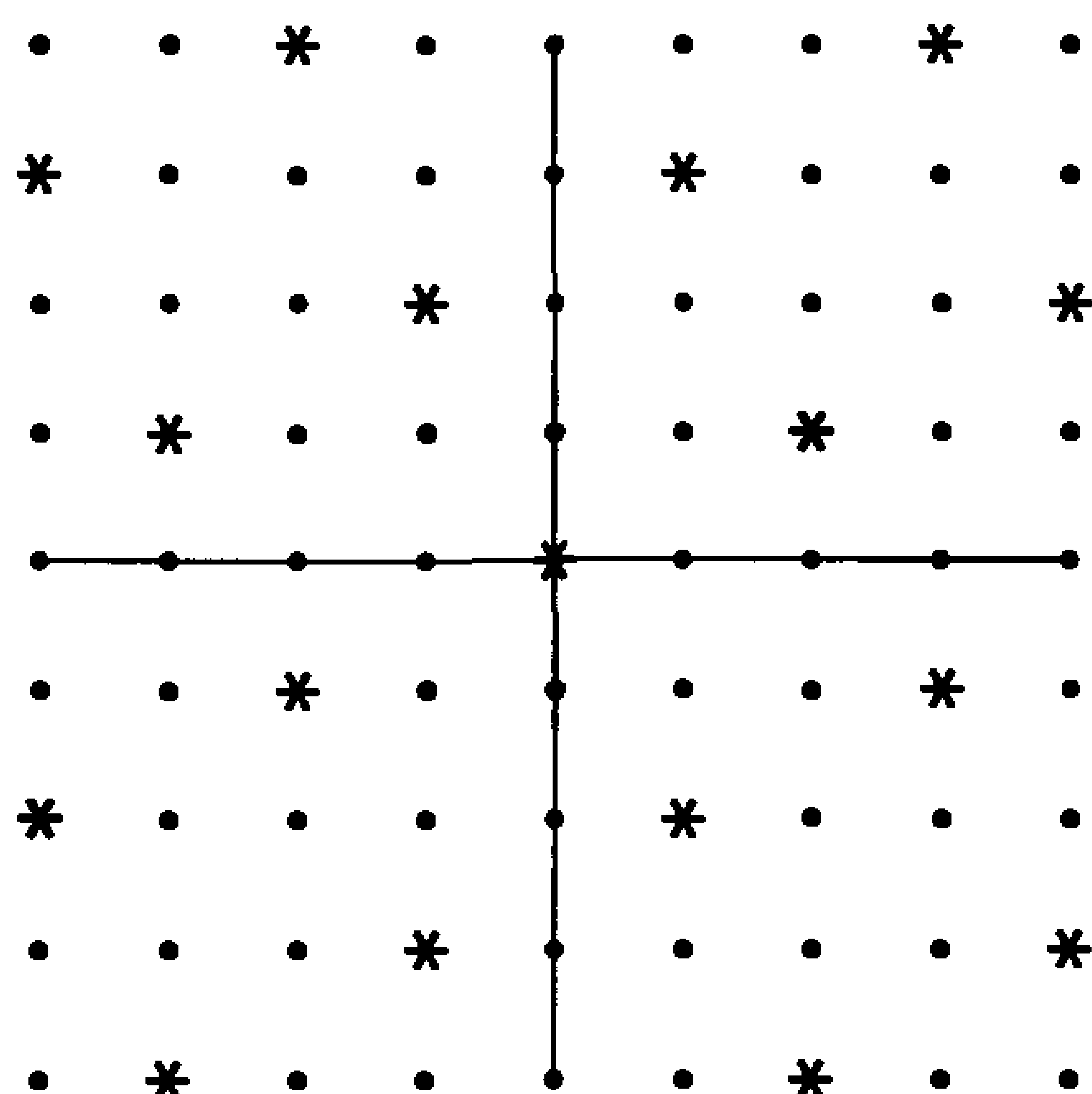
from the set of nonzero elements of R to the nonnegative integers. An integral domain R is a *Euclidean domain* if there is a size function σ on R such that the division algorithm holds:

(2.17) *Let $a, b \in R$ and suppose that $a \neq 0$. There are elements $q, r \in R$ such that $b = aq + r$, and either $r = 0$ or $\sigma(r) < \sigma(a)$.*

We do not require the elements q, r to be uniquely determined by a and b.

(2.18) **Proposition.** The rings $\mathbb{Z}$, $F[x]$, and $\mathbb{Z}[i]$ are Euclidean domains. □

The ring of integers and the polynomial ring have already been discussed. Let us show that the ring of Gauss integers is a Euclidean domain, with size function the function $\sigma = |\ \ |^2$. The elements of $\mathbb{Z}[i]$ form a square lattice in the complex plane, and the multiples of a given element a form a *similar lattice*, the ideal $(a) = Ra$. If we write $a = re^{i\theta}$, then (a) is obtained by rotating through the angle θ followed by stretching by the factor $r = |a|$:



(2.19) **Figure.** $*$ = ideal (a), $R = \mathbb{Z}[i]$

It is clear that for every complex number b, there is at least one point of the lattice (a) whose square distance from b is $\leq \frac{1}{2}|a|^2$. Let that point be aq, and set $r = b - aq$. Then $|r|^2 \leq \frac{1}{2}|a|^2 < |a|^2$, as required. Note that since there may be more than one choice for the element aq, this division with remainder is not unique.

We could also proceed algebraically. We divide the complex number b by a: $b = aw$, where $w = x + yi$ is a complex number, not necessarily a Gauss integer. Then we choose the nearest Gauss integer point (m, n) to (x, y), writing $x = m + x_0$, $y = n + y_0$, where m, n are integers and x_0, y_0 are real numbers such that $-\frac{1}{2} \leq x_0, y_0 < \frac{1}{2}$. Then $(m + ni)a$ is the required point of Ra. For, $|x_0 + y_0 i|^2 < \frac{1}{2}$ and $|b - (m + ni)a|^2 = |a(x_0 + y_0 i)|^2 < \frac{1}{2}|a|^2$.

One can copy the discussion of factorization of integers with minor changes to prove this proposition:

(2.20) **Proposition.** A Euclidean domain is a principal ideal domain, and hence it is a unique factorization domain. □

(2.21) **Corollary.** The rings $\mathbb{Z}$, $\mathbb{Z}[i]$, and $F[x]$ (F a field) are principal ideal domains and unique factorization domains. □

In the ring $\mathbb{Z}[i]$ of Gauss integers, the element 3 is irreducible, hence prime, but 2 and 5 are not irreducible because

(2.22)
$$2 = (1 + i)(1 - i) \quad \text{and} \quad 5 = (2 + i)((2 - i).$$

These are the prime factorizations of 2 and 5 in $\mathbb{Z}[i]$.

There are four units in the ring $\mathbb{Z}[i]$, namely $\{\pm 1, \pm i\}$. So every nonzero element α of this ring has four associates, namely the elements $\pm\alpha, \pm i\alpha$. The associates of $2 + i$, for example are

$$2 + i, \quad -2 - i, \quad -1 + 2i, \quad 1 - 2i.$$

There is no really natural way to normalize primes in $\mathbb{Z}[i]$, though if pressed we would choose the unique associate lying in the first quadrant and not on the imaginary axis. It is better to accept the ambiguity of (2.5) here or else work with principal ideals.

3. GAUSS'S LEMMA

Theorem (1.5) applies to the ring $\mathbb{Q}[x]$ of polynomials with rational coefficients: Every polynomial $f(x) \in \mathbb{Q}[x]$ can be expressed uniquely in the form $cp_1 \cdots p_k$, where $c \in \mathbb{Q}$ and p_i are monic polynomials which are irreducible over $\mathbb{Q}$. Now suppose that a polynomial $f(x)$ has integer coefficients, $f(x) \in \mathbb{Z}[x]$, and that it factors in $\mathbb{Q}[x]$. Can it be factored without leaving $\mathbb{Z}[x]$? We are going to prove that it can, and that $\mathbb{Z}[x]$ is a unique factorization domain.

Here is an example of a prime factorization in $\mathbb{Z}[x]$:

$$6x^3 + 9x^2 + 9x + 3 = 3(2x + 1)(x^2 + x + 1).$$

As we see from this example, irreducible factorizations are slightly more complicated in $\mathbb{Z}[x]$ than in $\mathbb{Q}[x]$. First, the prime integers are irreducible elements of $\mathbb{Z}[x]$, so they may appear in the prime factorization of a polynomial. Second, the factor $2x + 1$ isn't monic. If we want to stay with integer coefficients, we can't ask for monic factors.

The integer factors of a polynomial $f(x) = a_n x^n + \cdots + a_0$ in $\mathbb{Z}[x]$ are common divisors of its coefficients $a_0, \ldots, a_n$. A polynomial $f(x)$ is called *primitive* if its coefficients $a_0, \ldots, a_n$ have no common integer factor except for the units ± 1 and if its highest coefficient a_n is positive.

(3.1) **Lemma.** Every nonzero polynomial $f(x) \in \mathbb{Q}[x]$ can be written as a product

$$f(x) = cf_0(x),$$

where c is a rational number and $f_0(x)$ is a primitive polynomial in $\mathbb{Z}[x]$. Moreover, this expression for f is unique. The polynomial f has integer coefficients if and only if c is an integer. If so, then $|c|$ is the greatest common divisor of the coefficients of f, and the sign of c is the sign of the leading coefficient of f.

The rational number c which appears in this lemma is called the *content* of $f(x)$. If f has integer coefficients, then the content divides f in $\mathbb{Z}[x]$. Also, f is primitive if and only if its content is 1.

Proof of the Lemma. To find f_0, we first multiply f by an integer to clear the denominators in its coefficients. This will give us a polynomial f_1 with integer coefficients. Then we factor out the greatest common divisor of the coefficients of f_1 and adjust the sign of the leading coefficient. The resulting polynomial f_0 is primitive, and $f = cf_0$ for some rational number c. This proves existence.

To prove uniqueness, suppose that $cf_0(x) = dg_0(x)$, where $c, d \in \mathbb{Q}$ and f_0, g_0 are primitive polynomials. We will show that $c = d$ and $f_0 = g_0$. Clearing denominators reduces us to the case that c and d are integers. Let $\{a_i\}, \{b_i\}$ denote the coefficients of f_0, g_0 respectively. Then $ca_i = db_i$ for all i. Since the greatest common divisor of $\{a_0, \ldots, a_n\}$ is 1, c is the greatest common divisor of $\{ca_0, \ldots, ca_n\}$. Similarly, d is the greatest common divisor of $\{db_0, \ldots, db_n\} = \{ca_0, \ldots, ca_n\}$. Hence $c = \pm d$ and $f_0 = \pm g_0$. Since f_0 and g_0 have positive leading coefficients, $f_0 = g_0$ and $c = d$. If f has integer coefficients, clearing of the denominator is not necessary; hence c is an integer, and up to sign it is the greatest common divisor of the coefficients, as stated. □

As we have already observed, the Substitution Principle gives us a homomorphism

$$(3.2) \qquad \mathbb{Z}[x] \longrightarrow \mathbb{F}_p[x],$$

where $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ is the field with p elements. This homomorphism sends a polynomial $f(x) = a_m x^m + \cdots + a_0$ to its residue $\overline{f}(x) = \overline{a}_m x^m + \cdots + \overline{a}_0$ modulo p. We will now use it to prove Gauss's Lemma.

(3.3) Theorem. *Gauss's Lemma:* A product of primitive polynomials in $\mathbb{Z}[x]$ is primitive.

Proof. Let the polynomials be f and g, and let $h = fg$ be their product. Since the leading coefficients of f and g are positive, the leading coefficient of h is, too. To show that h is primitive, we will show that no prime integer p divides all the coefficients of $h(x)$. This will show that the content of h is 1. Consider the homomorphism $\mathbb{Z}[x] \longrightarrow \mathbb{F}_p[x]$ defined above. We have to show that $\bar{h} \neq 0$. Since f is primitive, its coefficients are not all divisible by p. So $\bar{f} \neq 0$. Similarly, $\bar{g} \neq 0$. Since the polynomial ring $\mathbb{F}_p[x]$ is an integral domain, $\bar{h} = \bar{f}\bar{g} \neq 0$, as required. □

(3.4) Proposition.

(a) Let f, g be polynomials in $\mathbb{Q}[x]$, and let f_0, g_0 be the associated primitive polynomials in $\mathbb{Z}[x]$. If f divides g in $\mathbb{Q}[x]$, then f_0 divides g_0 in $\mathbb{Z}[x]$.

(b) Let f be a primitive polynomial in $\mathbb{Z}[x]$, and let g be any polynomial with integer coefficients. Suppose that f divides g in $\mathbb{Q}[x]$, say $g = fq$, with $q \in \mathbb{Q}[x]$. Then $q \in \mathbb{Z}[x]$, and hence f divides g in $\mathbb{Z}[x]$.

(c) Let f, g be polynomials in $\mathbb{Z}[x]$. If they have a common nonconstant factor in $\mathbb{Q}[x]$, then they have a common nonconstant factor in $\mathbb{Z}[x]$ too.

Proof. To prove (a), we may clear denominators so that f and g become primitive. Then (a) is a consequence of (b). To prove (b), we apply (3.1) in order to write the quotient in the form $q = cq_0$, where q_0 is primitive and $c \in \mathbb{Q}$. By Gauss's Lemma, fq_0 is primitive, and the equation $g = cfq_0$ shows that it is the primitive polynomial g_0 associated to g. Therefore $g = cg_0$ is the expression for g referred to in Lemma (3.1), and c is the content of g. Since $g \in \mathbb{Z}[x]$, it follows that $c \in \mathbb{Z}$, hence that $q \in \mathbb{Z}[x]$. Finally, to prove (c), suppose that f, g have a common factor h in $\mathbb{Q}[x]$. We may assume that h is primitive, and then by (b) h divides both f and g in $\mathbb{Z}[x]$. □

(3.5) Corollary. If a nonconstant polynomial f is irreducible in $\mathbb{Z}[x]$, then it is irreducible in $\mathbb{Q}[x]$. □

(3.6) Proposition. Let f be an integer polynomial with positive leading coefficient. Then f is irreducible in $\mathbb{Z}[x]$ if and only if either

(i) f is a prime integer, or

(ii) f is a primitive polynomial which is irreducible in $\mathbb{Q}[x]$.

Proof. Suppose that f is irreducible. As in Lemma (3.1), we may write $f = cf_0$, where f_0 is primitive. Since f is irreducible, this can not be a proper factorization. So either c or f_0 is 1. If $f_0 = 1$, then f is constant, and to be irreducible, a constant polynomial must be a prime integer. If $c = 1$, then f is primitive, and is irreducible in $\mathbb{Q}[x]$ by the previous corollary. The converse, that integer primes and primitive irreducible polynomials are irreducible elements of $\mathbb{Z}[x]$, is clear. □

(3.7) **Proposition.** Every irreducible element of $\mathbb{Z}[x]$ is a prime element.

Proof. Let f be irreducible, and suppose f divides gh, where $g, h \in \mathbb{Z}[x]$.

Case 1: $f = p$ is a prime integer. Write $g = cg_0$ and $h = dh_0$ as in (3.1). Then g_0h_0 is primitive, and hence some coefficient a of g_0h_0 is not divisible by p. But since p divides gh, the corresponding coefficient, which is cda, is divisible by p. Hence p divides c or d, so p divides g or h.

Case 2: f is a primitive polynomial which is irreducible in $\mathbb{Q}[x]$. By (2.11b), f is a prime element of $\mathbb{Q}[x]$. Hence f divides g or h in $\mathbb{Q}[x]$. By (3.4), f divides g or h in $\mathbb{Z}[x]$. $\square$

(3.8) **Theorem.** The polynomial ring $\mathbb{Z}[x]$ is a unique factorization domain. Every nonzero polynomial $f(x) \in \mathbb{Z}[x]$ which is not ± 1 can be written as a product

$$f(x) = \pm p_1 \cdots p_m \, q_1(x) \cdots q_n(x),$$

where the p_i are prime integers and the $q_i(x)$ are irreducible primitive polynomials. This expression is unique up to arrangement of the factors.

Existence of factorizations is easy to prove for $\mathbb{Z}[x]$, so this theorem follows from Propositions (3.7) and (2.8). $\square$

Now let R be any unique factorization domain, and let F be its field of fractions [Chapter 10 (6.5)]. Then $R[x]$ is a subring of $F[x]$, and the results of this section can be copied, replacing $\mathbb{Z}$ by R and $\mathbb{Q}$ by F throughout. The only change to be made is that instead of normalizing primitive polynomials it is better to allow ambiguity caused by unit factors, as in the previous section. The main results are these:

(3.9) **Theorem.** Let R be a unique factorization domain with field of fractions F.

(a) Let f, g be polynomials in $F[x]$, and let f_0, g_0 be the associated primitive polynomials in $R[x]$. If f divides g in $F[x]$, then f_0 divides g_0 in $R[x]$.

(b) Let f be a primitive polynomial in $R[x]$, and let g be any polynomial in $R[x]$. Suppose that f divides g in $F[x]$, say $g = fq$, with $q \in F[x]$. Then $q \in R[x]$, and hence f divides g in $R[x]$.

(c) Let f, g be polynomials in $R[x]$. If they have a common nonconstant factor in $F[x]$, then they have a common nonconstant factor in $R[x]$ too.

(d) If a nonconstant polynomial f is irreducible in $R[x]$, then it is irreducible in $F[x]$.

(e) $R[x]$ is a unique factorization domain.

The proof of Theorem (3.9) follows the pattern established for the ring $\mathbb{Z}[x]$, and we omit it. $\square$

Since $R[x_1, \ldots, x_n] \approx R[x_1, \ldots, x_{n-1}][x_n]$, we obtain this corollary:

(3.10) **Corollary.** The polynomial rings $\mathbb{Z}[x_1, \ldots, x_n]$ and $F[x_1, \ldots, x_n]$, where F is a field, are unique factorization domains. $\square$

So the ring $\mathbb{C}[x,y]$ of complex polynomials in two variables is a unique factorization domain. In contrast to the case of one variable, however, where every complex polynomial is a product of linear ones, complex polynomials in two variables are often irreducible, and hence prime.

The irreducibility of a polynomial $f(x,y)$ can sometimes be proved by studying the locus $W = \{f(x,y) = 0\}$ in $\mathbb{C}^2$. Suppose that f factors, say

$$f(x,y) = g(x,y)h(x,y),$$

where g, h are nonconstant polynomials. Then $f(x,y) = 0$ if and only if one of the two equations $g(x,y) = 0$ or $h(x,y) = 0$ holds. So if we let $U = \{g(x,y) = 0\}$, $V = \{h(x,y) = 0\}$ denote these two varieties in $\mathbb{C}^2$, then

$$W = U \cup V.$$

It may be possible to see geometrically that W has no such decomposition.

For example, we can use this method to show that the polynomial

$$f(x,y) = x^2 + y^2 - 1$$

is irreducible. Since the total degree of f is 2, any proper factor of f has to be linear, of the form $g(x,y) = ax + by + c$. And the solutions to a linear equation lie on a line, whereas $\{f = 0\}$ is a circle. Of course when we speak of lines and circles, we are actually talking about the real loci in $\mathbb{R}^2$. So this reasoning shows that f is irreducible in $\mathbb{R}[x,y]$. But in fact, the real locus of a circle has enough points to show irreducibility in $\mathbb{C}[x,y]$ too. Suppose that $f = gh$ in $\mathbb{C}[x,y]$, where g and h are linear as before. Then every point of the real circle $x^2 + y^2 - 1 = 0$ lies on one of the complex loci U, V. So at least one of these loci contains two real points. There is exactly one complex line (a *line* being the locus of solutions of a linear equation $ax + by + c = 0$) which passes through two given points, and if these points are real, the linear equation defining the line is also real, up to a constant factor. This is proved by writing down the equation of a line through two points explicitly. So if f has a linear factor, then it has a real one. But the circle does not contain a line.

One can also prove that $x^2 + y^2 - 1$ is irreducible algebraically, using the method of undetermined coefficients (see Section 4, exercise 17).

4. EXPLICIT FACTORIZATION OF POLYNOMIALS

We now pose the problem of determining the factors of a given integer polynomial

$$f(x) = a_n x^n + \cdots + a_j x + a_0. \tag{4.1}$$

What we want are the irreducible factors in $\mathbb{Q}[x]$, and by (3.5) this amounts to determining the irreducible factors in $\mathbb{Z}[x]$. Linear factors can be found fairly easily. If $b_1 x + b_0$ divides $f(x)$, then b_1 divides a_n and b_0 divides a_0. There are finitely many integers which divide a_n and a_0, so we can try all possibilities. In each case, we carry out the division and determine whether the remainder is zero. Or we may substitute the rational number $r = -b_0/b_1$ into $f(x)$ to see if it is a root.

Though things are not so clear for factors of higher degree, Kronecker showed that the factors can be determined with a finite number of computations. His method is based on the Lagrange interpolation formula. Unfortunately this method requires too many steps to be practical except for factors of low degree, and a lot of work has been done on the problem of efficient computation. One of the most useful methods is computation modulo p, using the homomorphism $\mathbb{Z}[x] \longrightarrow \mathbb{F}_p[x]$. If our polynomial $f(x)$ factors in $\mathbb{Z}[x]$: $f = gh$, then its residue $\bar{f}(x)$ modulo p also factors: $\bar{f} = \bar{g}\bar{h}$. And since there are only finitely many polynomials of each degree in $\mathbb{F}_p[x]$, all factorizations there can be carried out in finitely many steps.

(4.2) **Proposition.** Let $f(x) = a_n x^n + \cdots + a_0 \in \mathbb{Z}[x]$ be an integer polynomial, and let p be a prime integer which does not divide a_n. If the residue $\bar{f}$ of f modulo p is irreducible, then f is irreducible in $\mathbb{Q}[x]$.

Proof. This follows from an inspection of the homomorphism. We need the assumption that p does not divide a_n in order to rule out the possibility that a factor g of f could reduce to a constant in $\mathbb{F}_p[x]$. This assumption is preserved if we replace f by the associated primitive polynomial. So we may assume that f is primitive. Since p does not divide a_n, the degrees of f and $\bar{f}$ are equal. If f factors in $\mathbb{Q}[x]$, then it also factors in $\mathbb{Z}[x]$, by Corollary (3.5). Let $f = gh$ be a proper factorization in $\mathbb{Z}[x]$. Since f is primitive, g and h have positive degree. Since $\deg f = \deg \bar{f}$ and $\bar{f} = \bar{g}\bar{h}$, it follows that $\deg g = \deg \bar{g}$ and $\deg h = \deg \bar{h}$, hence that $\bar{f} = \bar{g}\bar{h}$ is a proper factorization, which shows that $\bar{f}$ is reducible. □

Suppose we suspect that a given polynomial $f(x) \in \mathbb{Z}[x]$ is irreducible. Then we can try reduction modulo p for a few low primes, $p = 2$ or 3 for instance, and hope that $\bar{f}$ turns out to be of the same degree and irreducible. If so, we will have proved that f is irreducible too. Note also that since $\mathbb{F}_p$ is a field, the results of Theorem (1.5) hold for the ring $\mathbb{F}_p[x]$.

Unfortunately, there exist integer polynomials which are irreducible, though they can be factored modulo p for every prime p. The polynomial $x^4 - 10x^2 + 1$ is an example. So the method of reduction modulo p will not always work. But it does work quite often.

The irreducible polynomials in $\mathbb{F}_p[x]$ can be found by the "sieve" method. The *sieve of Eratosthenes* is the name given to the following method of determining the primes less than a given number n. We list the integers from 2 to n. The first one, 2, is prime because any proper factor of 2 must be smaller than 2, and there is no smaller integer on the list. We make a note of the fact that 2 is prime, and then we cross out the multiples of 2 from our list. Except for 2 itself, they are not prime. The first integer which is left, 3, is a prime because it isn't divisible by any smaller prime. We note that 3 is a prime and then cross out the multiples of 3 from our list. Again, the smallest remaining integer, 5, is a prime, and so on.

2 3 ~~4~~ 5 ~~6~~ 7 ~~8~~ ~~9~~ ~~10~~ 11 ~~12~~ 13 ~~14~~ ~~15~~ ~~16~~ 17 ~~18~~ 19 ....

This method will also determine the irreducible polynomials in $\mathbb{F}_p[x]$. We list all polynomials, degree by degree, and then cross out products. For example, the

linear polynomials in $\mathbb{F}_2[x]$ are x and $x + 1$. They are irreducible. The polynomials of degree 2 are x^2, $x^2 + x$, $x^2 + 1$, and $x^2 + x + 1$. The first three are divisible by x or by $x + 1$, so the last one is the only irreducible polynomial of degree 2 over $\mathbb{F}_2$.

(4.3) *The irreducible polynomials of degree* ≤ 4 *over* $\mathbb{F}_2$:

$$x, \quad x + 1; \; x^2 + x + 1; \; x^3 + x^2 + 1, \quad x^3 + x + 1;$$

$$x^4 + x^3 + 1, \quad x^4 + x + 1, \quad x^4 + x^3 + x^2 + x + 1.$$

By trying the polynomials on this list, we can factor all polynomials of degree 9 or less in $\mathbb{F}_2[x]$.

As a sample application of 4.2, the polynomial $x^4 - 6x^3 + 12x^2 - 3x + 9$ is irreducible in $\mathbb{Q}[x]$, because its residue in $\mathbb{F}_2[x]$ is $x^4 + x + 1$.

(4.4) *The monic irreducible polynomials of degree* 2 *over* $\mathbb{F}_3$:

$$x^2 + 1, \quad x^2 + x - 1, \quad x^2 - x - 1.$$

Reduction modulo p may help describe the factorization of a polynomial even though the residue is reducible. Consider the polynomial $f(x) = x^3 + 6x + 3$ for instance. Reducing modulo 3, we obtain x^3. This doesn't look like a promising tool. However, suppose that $f(x)$ were reducible, say $(ax + b)(cx^2 + dx + e) = x^3 + 6x + 3$. Then the residue of $ax + b$ would have to divide x^3 in $\mathbb{F}_3[x]$, which would imply $b \equiv 0$ (modulo 3). Similarly, we could conclude $e \equiv 0$ (modulo 3). It is impossible to satisfy both of these conditions, because $be = 3$. Therefore no such factorization exists, and $f(x)$ is irreducible.

The principle at work in this example is called the Eisenstein Criterion.

(4.5) **Proposition.** *Eisenstein Criterion:* Let $f(x) = a_n x^n + \cdots + a_0 \in \mathbb{Z}[x]$ be an integer polynomial, and let p be a prime integer. Suppose that the coefficients of f satisfy the following conditions:

(i) p does not divide a_n;
(ii) p divides the other coefficients $a_{n-1}, \ldots, a_0$;
(iii) p^2 does not divide a_0.

Then f is irreducible in $\mathbb{Q}[x]$. If f is primitive, it is irreducible in $\mathbb{Z}[x]$.

For example, $x^4 + 50x^2 + 30x + 20$ is irreducible in $\mathbb{Q}[x]$ and in $\mathbb{Z}[x]$.

Proof of the Eisenstein Criterion. Assume that the conditions are met for f. Let $\overline{f}$ denote the residue modulo p. The hypotheses (i) and (ii) imply that $\overline{f} = \overline{a}_n x^n$ and that $\overline{a}_n \neq 0$. If f is reducible in $\mathbb{Q}[x]$, then it will factor in $\mathbb{Z}[x]$ into factors of positive degree, say $f = gh$. Then $\overline{g}$ and $\overline{h}$ divide $\overline{a}_n x^n$, and hence each of these polynomials is a monomial. Therefore all coefficients of g and of h, except the highest ones, are divisible by p. Let the constant coefficients of g, h be b_0, c_0. Then the constant coefficient of f is $a_0 = b_0 c_0$. Since p divides b_0 and c_0, it follows that p^2 divides a_0, which contradicts (iii). This shows that f is irreducible. The last assertion follows from (3.6). □

One of the most important applications of the Eisenstein Criterion is to prove the irreducibility of the *cyclotomic polynomial* $x^{p-1} + x^{p-2} + \cdots + x + 1$, whose roots are the pth roots of unity, the powers of $\zeta = e^{2\pi i}/p$:

(4.6) **Corollary.** Let p be a prime. The polynomial $f(x) = x^{p-1} + x^{p-2} + \cdots + x + 1$ is irreducible in $\mathbb{Q}[x]$.

Proof. We note that $(x - 1)f(x) = x^p - 1$. Next, we make the substitution $x = y + 1$ into this product, obtaining

$$yf(y + 1) = (y + 1)^p - 1 = y^p + \binom{p}{1}y^{p-1} + \binom{p}{2}y^{p-2} + \cdots + \binom{p}{p-1}y.$$

We have $\binom{p}{i} = p(p - 1)\cdots(p - i - 1)/i!$. If $i < p$, then the prime p isn't a factor of $i!$, so $i!$ divides the product $(p - 1)\cdots(p - i + 1)$ of the remaining terms in the numerator of the integer $\binom{p}{i}$. This implies that $\binom{p}{i}$ is divisible by p. Dividing the expansion of $yf(y + 1)$ by y shows that $f(y + 1)$ satisfies the conditions of the Eisenstein Criterion, hence that it is an irreducible polynomial. It follows that $f(x)$ is irreducible too. □

It is instructive to examine the statement analogous to the Eisenstein Criterion when the ring of integers is replaced by the polynomial ring $\mathbb{C}[t]$. Then $\mathbb{Z}[x]$ gets replaced by $\mathbb{C}[t][x] \approx \mathbb{C}[t, x]$, the polynomial ring in two variables.

(4.7) **Proposition.** Let $f(t, x)$ be an element of $\mathbb{C}[t, x]$, written as a polynomial in x whose coefficients are polynomials in t: $f(t, x) = a_n(t)x^n + \cdots + a_1(t)x + a_0(t)$. Suppose that

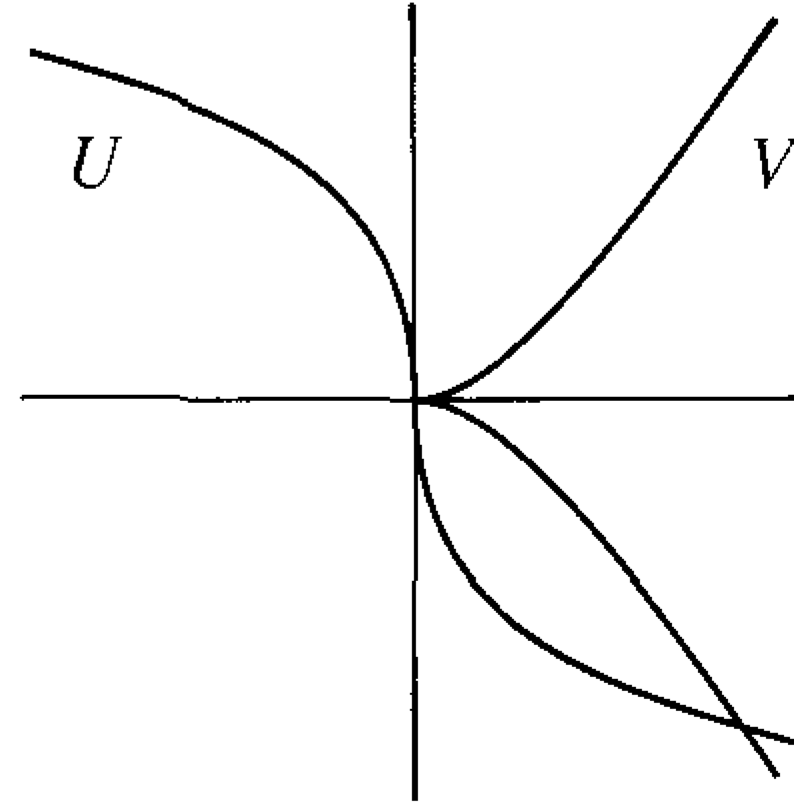
(i) t does not divide $a_n(t)$,
(ii) t divides $a_{n-1}(t), \ldots, a_0(t)$,
(iii) t^2 does not divide $a_0(t)$.

Then $f(t, x)$ is irreducible in the ring $\mathbb{C}(t)[x]$. If f is primitive, meaning that it has no factor which is a polynomial in t alone, then f is irreducible in $\mathbb{C}[t, x]$.

This can be proved exactly as we proved (4.5), replacing $\mathbb{F}_p[x]$ by $\mathbb{C}[x] = \mathbb{C}[t, x]/(t)$. But let us examine the geometry of this situation by considering the locus $W = \{f(t, x) = 0\}$ in complex 2-space. Conditions (i) and (ii) of (4.7) imply that $f(0, x) = cx^n$, where $c = a_n(0) \neq 0$. Consequently the only solution of $f(t, x) = 0$ with $t = 0$ is $t = x = 0$, so the variety W meets the x-axis $\{t = 0\}$ only at the origin.

Suppose that $f(t, x)$ is reducible: $f(t, x) = g(t, x)h(t, x)$. Then W is the union of the two varieties $U = \{g = 0\}$ and $V = \{h = 0\}$. Also, $cx^n = f(0, x) = g(0, x)h(0, x)$. Hence $g(0, x)$ is a constant times x^r, and $h(0, x)$ is a constant times x^{n-r}, where r is the degree of g in the variable x. Therefore g and h both vanish at

the origin. It follows that the origin is a *singular point* of W, meaning that the partial derivatives $\partial f/\partial x$ and $\partial f/\partial t$ both vanish at $(0,0)$. This is checked by differentiating the product gh. On the other hand, $\partial f/\partial t\,(0,0) = da_0/dt\,(0)$, and this is the linear coefficient of $a_0(t)$. If it vanishes, t^2 divides $a_0(t)$, contrary to (4.7iii). □



5. PRIMES IN THE RING OF GAUSS INTEGERS

We have seen that the ring of Gauss integers is a Euclidean domain. Its units are $\{\pm 1, \pm i\}$, and every element which is not zero and not a unit is a product of prime elements. In this section we will study these prime elements, called *Gauss primes*, and their relation to prime integers. We looked at some examples in Section 2, where we saw that the prime integer 5 factors in $\mathbb{Z}[i]$: $5 = (2+i)(2-i)$, while 3 does not factor; 3 is a Gauss prime. Remember that since there are four units, there are four associate factorizations of the integer 5 which we consider equivalent:

$$(2+i)(2-i) = (-2-i)(-2+i) = (1-2i)(1+2i) = (-1+2i)(-1-2i).$$

We will now show that the examples 3 and 5 exhibit the two ways that prime integers can factor in the ring $\mathbb{Z}[i]$. The story is summed up in this theorem:

(5.1) **Theorem.**

(a) Let p be a prime integer. Then either p is a Gauss prime, or else it is the product of two complex conjugate Gauss primes: $p = \pi\overline{\pi}$.

(b) Let π be a Gauss prime. Then either $\pi\overline{\pi}$ is a prime integer, or else it is the square of a prime integer.

(c) The prime integers which are Gauss primes are those congruent to 3 modulo 4; that is, $p = 3, 7, 11, 19, \ldots$.

(d) Let p be a prime integer. The following are equivalent:

(i) p is a product of two complex conjugate Gauss primes.
(ii) p is the sum of two integer squares: $p = a^2 + b^2$, with $a, b \in \mathbb{Z}$.
(iii) The congruence $x^2 \equiv -1$ (modulo p) has an integer solution.
(iv) $p \equiv 1$ (modulo 4), or $p = 2$; that is, $p = 2, 5, 13, 17, \ldots$.

It will take some time to prove all parts of this theorem.

The following lemma follows directly from the definition of a Gauss integer:

(5.2) **Lemma.** A Gauss integer which is a real number is an ordinary integer. An ordinary integer d divides another integer a in $\mathbb{Z}[i]$ if and only if d divides a in $\mathbb{Z}$. Moreover, d divides a Gauss integer $a + bi$ if and only if d divides both a and b.

Now to prove part (a) of the theorem, let p be an integer prime. Then p is not a unit in the ring $\mathbb{Z}[i]$. Hence it has a Gauss prime divisor, say $\pi = a + bi$, where $a, b \in \mathbb{Z}$. The complex conjugate $\overline{\pi} = a - bi$ also divides p because $p = \overline{p}$, so $\pi\overline{\pi} = a^2 + b^2$ divides p^2 in the ring of Gauss integers. Being an integer, $\pi\overline{\pi}$ is an integer divisor of p^2. There are two possibilities: π may be an associate of p. In this case, p is a Gauss prime. Otherwise π is a proper divisor of p in the ring of Gauss integers, and then $\pi\overline{\pi}$ is a proper divisor of p^2 in the ring $\mathbb{Z}$. Since $\pi\overline{\pi}$ is a positive integer, $\pi\overline{\pi} = p$ in this case.

We can turn this argument around to prove (b). Let π be a Gauss prime. Then $\pi\overline{\pi}$ is a positive integer, say $\pi\overline{\pi} = n$. We factor n into primes in the ring of integers. This factorization will also be a factorization in the Gauss integers, though not necessarily a prime factorization. Since π is a Gauss prime which divides n in $\mathbb{Z}[i]$, it divides one of the integer prime factors of n. Thus π divides an integer prime p. Then $\pi\overline{\pi}$ is an integer divisor of p^2, hence $\pi\overline{\pi} = p$ or p^2.

Note that part (c) of Theorem (5.1) is a formal consequence of (a) and of the equivalence of conditions (d)(i) and (d)(iv). So we need not consider part (c) further, and we now turn to the proof of part (d). It is easy to see that (i) and (ii) of part (d) are equivalent: Suppose that $p = \pi\overline{\pi}$ for some Gauss prime $\pi = a + bi$. Then $p = \pi\overline{\pi} = (a + bi)(a - bi) = a^2 + b^2$, so p is a sum of two integer squares. Conversely, if $p = a^2 + b^2$, then $p = (a + bi)(a - bi)$ provides a factorization of p in the ring of Gauss integers, which is a prime factorization because of (a).

The equivalence of (d)(i) and (d)(iii) of Theorem (5.1) is harder to prove. To do so, we go back to the formal construction of the Gauss integers. The ring $\mathbb{Z}[i]$ is obtained from the ring $\mathbb{Z}$ by adjoining an element i with the relation $i^2 + 1 = 0$. So there is an isomorphism

$$\mathbb{Z}[x]/(x^2 + 1) \xrightarrow{\sim} \mathbb{Z}[i]. \tag{5.3}$$

Let (p) denote the principal ideal generated by a prime integer p in the ring of Gauss integers. Its elements are the Gauss integers $a + bi$ such that a and b are both divisible by p. Denote by R' the quotient ring $\mathbb{Z}[i]/(p)$. Then R' can also be thought of as the ring obtained by introducing the two relations

$$x^2 + 1 = 0 \quad \text{and} \quad p = 0 \tag{5.4}$$

into the polynomial ring $\mathbb{Z}[x]$. So we have an isomorphism

$$\mathbb{Z}[x]/(x^2 + 1, p) \xrightarrow{\sim} \mathbb{Z}[i]/(p) = R', \tag{5.5}$$

where $(x^2 + 1, p)$ denotes the ideal of $\mathbb{Z}[x]$ generated by the two elements.

(5.6) **Lemma.** Let p be a prime integer. The following statements are equivalent:

(i) p is a Gauss prime;

(ii) the ring $R' = \mathbb{Z}[i]/(p)$ is a field;
(iii) $x^2 + 1$ is an irreducible polynomial in the ring $\mathbb{F}_p[x]$.

Proof. The equivalence of the first two statements follows from Proposition (2.14). What we are really after is the equivalence of (i) and (iii), and at first glance, these two statements do not seem to be related at all. It was in order to obtain this equivalence that we introduced the auxiliary ring R'. The proof is based on the following elementary but remarkably useful observation, which follows from the Third Isomorphism Theorem [Chapter 10 (4.3b)]:

(5.7) *To construct the ring R', it does not matter which of the two relations (5.4) is introduced into the ring $\mathbb{Z}[x]$ first.*

So let us reverse the order and begin by killing the element p. The Substitution Principle tells us what we will get. The kernel of the homomorphism $\mathbb{Z}[x] \longrightarrow \mathbb{F}_p[x]$ is precisely the ideal $p\mathbb{Z}[x]$. Since this map is surjective, it induces an isomorphism

$$\mathbb{Z}[x]/p\mathbb{Z}[x] \xrightarrow{\sim} \mathbb{F}_p[x].$$

We now introduce our other relation $x^2 + 1 = 0$ into this ring, interpreting the coefficients of this polynomial as elements of $\mathbb{F}_p$. The result is an isomorphism

$$\mathbb{F}_p[x]/(x^2 + 1) \xrightarrow{\sim} R'. \tag{5.8}$$

Proposition (2.14), applied to the ring $\mathbb{F}_p[x]$, shows that R' is a field if and only if $x^2 + 1$ is irreducible in $\mathbb{F}_p[x]$. □

We can now prove the equivalence of conditions (d)(i) and (d)(iii) of (5.1). We know by Lemma (5.6) that p is a Gauss prime if and only if $x^2 + 1$ is an irreducible polynomial in the ring $\mathbb{F}_p[x]$. Since it is a quadratic polynomial, $x^2 + 1$ is reducible if it has a root in $\mathbb{F}_p$ and irreducible if it has no root. Also, the residue of an integer a (modulo p) is a root of $x^2 + 1$ if and only if $a^2 \equiv -1$ (modulo p). Thus the congruence $x^2 \equiv -1$ (modulo p) has a solution if and only if $x^2 + 1$ is reducible modulo p, which happens if and only if p is not a Gauss prime. The equivalence of (i) and (iii) follows.

It remains to prove the equivalence of condition (iv) of part (d) with one of the other conditions. We will show its equivalence with condition (iii). The congruence $x^2 \equiv -1$ (modulo 2) does have the solution $x = 1$, so it is sufficient to look at the other primes, that is, at the odd primes. The following lemma does the job:

(5.9) **Lemma.** Let p be an odd prime, and let $\overline{a}$ denote the residue of an integer a modulo p.

(a) The integer a solves the congruence $x^2 \equiv -1$ (modulo p) if and only if its residue $\overline{a}$ is an element of order 4 in the multiplicative group of the field $\mathbb{F}_p$.
(b) The multiplicative group $\mathbb{F}_p^\times$ contains an element of order 4 if and only if $p \equiv 1$ (modulo 4).

Proof. There is exactly one element of order 2 in $\mathbb{F}_p^{\times}$, namely the residue of -1. This is because an element of order 2 is a root of the polynomial $x^2 - 1$, and we know the roots of this polynomial: They are ± 1 in any field [see (1.7)]. If a residue $\overline{a}$ has order 4 in $\mathbb{F}_p^{\times}$, then $\overline{a}^2$ has order 2; hence $\overline{a}^2 = -1$, which means $a^2 \equiv -1$ (modulo p). Conversely, if $a^2 \equiv -1$ (modulo p), then $\overline{a}$ has order 4 in $\mathbb{F}_p^{\times}$. This proves part (a) of the lemma.

Now the order of the group $\mathbb{F}_p^{\times}$ is $p - 1$. So if this group contains an element of order 4, then $p - 1$ is divisible by 4, or equivalently $p \equiv 1$ (modulo 4). Conversely, suppose that $p - 1$ is divisible by 4, and let H be the Sylow-2 subgroup of $\mathbb{F}_p^{\times}$, whose order is the largest power 2^r of 2 which divides $p - 1$. Since 4 divides $p - 1$, the order of H is at least 4, so there is an element $\overline{a}$ in H different from ± 1. This element does not have order 2, nor does it have order 1. But since H is a 2-group, the order of $\overline{a}$ is a power of 2. So some power of $\overline{a}$ has order exactly 4.

This completes the proof of Theorem (5.1). □

6. ALGEBRAIC INTEGERS

In the next sections we are going to study factorization of algebraic numbers in a simple but important case, that of quadratic imaginary integers. The ring of Gauss integers is our model here. It was in order to extend the properties of factorization of ordinary integers to algebraic numbers that ideals were first introduced, and the extension is very beautiful.

In contrast to most of the topics we have studied, the arithmetic of quadratic number fields is not of universal importance. It has many applications to arithmetic, but not so many in other areas of mathematics. Our reason for including this topic, aside from its elegance, is its historical importance. Many of our algebraic tools were first developed in order to extend arithmetic properties of the integers to algebraic numbers.

A typical application of algebraic numbers to arithmetic is to the problem of determining integer points on an ellipse such as

$$x^2 + 5y^2 = p, \tag{6.1}$$

where for simplicity we assume that p is a prime. To determine integer points on the circle $x^2 + y^2 = p$, we may begin by factoring the left side, obtaining $(x + iy)(x - iy) = p$, and then use arithmetic in the Gauss integers to analyze the factorization. We did this in our proof of Theorem (5.1). The analogous procedure for equation (6.1) leads to

$$(x + \sqrt{-5}y)(x - \sqrt{-5}y) = p,$$

so we may attempt an analysis in the ring $\mathbb{Z}[\sqrt{-5}]$. However, as we have seen, factorization is not unique in this ring. We will have some trouble.

Another example is the famous Fermat Equation

$$x^3 + y^3 = z^3. \tag{6.2}$$

It was proved by Euler that this equation has no integer solutions, except for the trivial solutions in which one of the variables is zero. To analyze it, we may bring y^3 to the other side and factor, obtaining

$$x^3 = (z - y)(z - \zeta y)(z - \bar{\zeta} y), \tag{6.3}$$

where

$$\zeta = \tfrac{1}{2}(-1 + \sqrt{-3}) = e^{2\pi i/3} \tag{6.4}$$

is a complex cube root of 1. One can then analyze this equation using arithmetic in the ring $\mathbb{Z}[\zeta]$. This ring happens to be a Euclidean domain, so unique factorization is available. Unfortunately, the proof that (6.2) has no nontrivial solution is fairly complicated, so we will not give it.

Problems of this type, which ask for integer solutions of polynomial equations, are called *Diophantine problems*. We will analyze a few of them in Section 12, when the necessary tools have been assembled.

A complex number α is called algebraic if it is the root of a nonzero polynomial $f(x)$ with rational coefficients (Chapter 10, Section 1). We can, of course, clear denominators in the coefficients of the polynomial $f(x)$. So if α is an algebraic number, then it is also the root of a polynomial with integer coefficients. The number α is called an *algebraic integer* if it is the root of a *monic* polynomial with integer coefficients, a polynomial of the form

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0, \quad \text{with } a_i \in \mathbb{Z}. \tag{6.5}$$

Thus the cube root of unity ζ, being a root of the polynomial $x^3 - 1$, is an algebraic integer.

Let α be an algebraic number. The set of all polynomials in $\mathbb{Q}[x]$ which have α as a root is the kernel of the substitution homomorphism

$$\mathbb{Q}[x] \longrightarrow \mathbb{C}, \quad \text{defined by } f(x) \rightsquigarrow f(\alpha).$$

So it is a principal ideal, generated by an irreducible element $f(x)$ of the polynomial ring which is called the *irreducible polynomial for α over $\mathbb{Q}$*. (Why is f irreducible?) It is the polynomial of lowest degree having α as a root and is unique up to a constant factor. The degree of the irreducible polynomial for α is also called the *degree* of α over $\mathbb{Q}$.

We may choose this irreducible polynomial $f(x)$ for α to be a *primitive* polynomial in $\mathbb{Z}[x]$. Then $f(x)$ also generates the ideal of $\mathbb{Z}[x]$ of all integer polynomials having α as a root.

(6.6) **Proposition.** The kernel of the map $\mathbb{Z}[x] \longrightarrow \mathbb{C}$ sending $x \rightsquigarrow \alpha$ is the principal ideal of $\mathbb{Z}[x]$ generated by the primitive irreducible polynomial for α.

Proof. Let $f(x)$ be the primitive irreducible polynomial for α. If $g \in \mathbb{Z}[x]$ has α as a root, then f divides g in $\mathbb{Q}[x]$, and hence f divides g in $\mathbb{Z}[x]$ too, by (3.4). So g is in the principal ideal of $\mathbb{Z}[x]$ generated by f. □

Note that the leading coefficient of a polynomial $f(x)$ divides the leading coefficient of any multiple in $\mathbb{Z}[x]$. So it follows from Proposition (6.6) that if the primitive irreducible polynomial $f(x)$ for α is *not* monic, then α is not the root of any monic integer polynomial.

(6.7) **Proposition.** An algebraic number α is an algebraic integer if and only if the primitive irreducible polynomial for α is monic. Equivalently, α is an algebraic integer if and only if the monic irreducible polynomial for α in $\mathbb{Q}[x]$ has integer coefficients. □

The primitive irreducible polynomial for the cube root of unity ζ is $x^2 + x + 1$.

(6.8) **Corollary.** A rational number r is an algebraic integer if and only if it is an ordinary integer.

For, the monic irreducible polynomial over $\mathbb{Q}$ of a rational number r is $x - r$. □

Proposition (6.7) can be used to decide whether or not an algebraic number is an algebraic integer, provided that we can compute its irreducible polynomial. For example, $\alpha = \frac{1}{2}(1 + \sqrt{2})$ is a root of $4x^2 - 4x - 1$. This is the primitive irreducible polynomial for α. Hence α is not an algebraic integer.

The concept of algebraic integer was one of the most important discoveries of number theory. It is not easy to explain quickly why it is the right definition to use, but roughly speaking, we can think of the leading coefficient of the primitive irreducible polynomials $f(x)$ for α as a "denominator." If α is the root of an integer polynomial $f(x) = dx^n + a_{n-1}x^{n-1} + \cdots + a_0$, then $d\alpha$ is an algebraic integer, because it is a root of the monic integer polynomial

$$x^n + a_{n-1}x^{n-1} + da_{n-2}x^{n-2} + \cdots + d^{n-2}a_1x + d^{n-1}a_0. \tag{6.9}$$

Thus we can "clear the denominator" in any algebraic number α by multiplying it with a suitable integer to get an algebraic integer. The leading coefficient is, however, not a precise denominator. Thus if $\alpha = \frac{1}{2}(1 + \sqrt{2})$, then 2α is an algebraic integer, while the leading coefficient of its primitive irreducible polynomial is 4.

In another direction, the example of the algebraic integer $\zeta = \frac{1}{2}(-1 + \sqrt{-3})$ shows that we must not jump to conclusions just because some expression for an algebraic number has denominators.

Explicit computation with algebraic integers is not very easy. It is a fact that they form a subring of $\mathbb{C}$, that is, that sums and products of algebraic integers are algebraic integers, but this isn't obvious. Rather than develop a general theory, we will work out the case of quadratic extensions explicitly.

A *quadratic number field* $F = \mathbb{Q}[\sqrt{d}]$ consists of all complex numbers

$$a + b\sqrt{d}, \quad \text{with} \quad a,b \in \mathbb{Q}, \tag{6.10}$$

where d is a fixed integer, positive or negative, which is not a rational square. The notation $\sqrt{d}$ will stand for the positive square root if $d > 0$ and for the positive

imaginary square root if $d < 0$. If d has a square integer factor, we can pull it out of the radical and put it into b without changing the field. Therefore it is customary to assume that d is *square free,* meaning that $d = \pm p_1 \cdots p_r$ where the p_i are distinct primes, or that $d = -1$. So the values we take are

$$d = -1, \pm 2, \pm 3, \pm 5, \pm 6, \pm 7, \pm 10, \ldots .$$

The field F is called a *real quadratic number field* if $d > 0$, or an *imaginary quadratic number field* if $d < 0$.

We will now compute the algebraic integers in F. The computation for a special value of d is no simpler than the general case. Nevertheless, you may wish to substitute a value such as $d = 5$ when going over this computation. We set

$$\delta = \sqrt{d}. \tag{6.11}$$

When d is negative, δ is purely imaginary. Let

$$\alpha = a + b\delta$$

be any element of F which is not in $\mathbb{Q}$, that is, such that $b \neq 0$. Then $\alpha' = a - b\delta$ is also in F. If d is negative, α' is the complex conjugate of α. Note that α is a root of the polynomial

$$(x - \alpha)(x - \alpha') = x^2 - (\alpha+\alpha')x + \alpha\alpha' = x^2 - 2ax + (a^2-b^2d). \tag{6.12}$$

This polynomial has the rational coefficients $-2a$ and $a^2 - b^2d$. Since α is not a rational number, it is not the root of a linear polynomial. So (6.12) is irreducible and is therefore the monic irreducible polynomial for α over $\mathbb{Q}$. According to (6.7), α is an algebraic integer if and only if (6.12) has integer coefficients. Thus we have the following corollary:

(6.13) **Corollary.** $\alpha = a + b\delta$ is an algebraic integer if and only if $2a$ and $a^2 - b^2d$ are integers. □

This corollary also holds when $b = 0$, because if a^2 is an integer, then so is a. If we like, we can use the conditions of the corollary as a definition of the integers in F.

The possibilities for a and b depend on the congruence class of d modulo 4. Note that since d is assumed to be square free, the case $d \equiv 0$ (modulo 4) has been ruled out, so $d \equiv 1, 2$, or 3 (modulo 4).

(6.14) **Proposition.** The algebraic integers in the quadratic field $F = \mathbb{Q}[\sqrt{d}]$ have the form $\alpha = a + b\delta$, where:

(a) If $d \equiv 2$ or 3 (modulo 4), then a and b are integers.
(b) If $d \equiv 1$ (modulo 4), then either $a, b \in \mathbb{Z}$ or $a, b \in \mathbb{Z} + \frac{1}{2}$.

The cube root of unity $\zeta = \frac{1}{2}(-1 + \sqrt{-3})$ is an example of an algebraic integer of the second type. On the other hand, since $-1 \equiv 3$ (modulo 4), the integers in the field $\mathbb{Q}[i]$ are just the Gauss integers.

Proof of the Proposition. Since the coefficients of the irreducible polynomial (6.12) for α are $2a$ and $a^2 - b^2d$, α is certainly an algebraic integer if a and b are integers. Assume that $d \equiv 1$ (modulo 4) and that $a, b \in \mathbb{Z} + \frac{1}{2}$. (We say that they are *half integers.*) Then $2a \in \mathbb{Z}$. To show that $a^2 - b^2d \in \mathbb{Z}$, we write $a = \frac{1}{2}m$, $b = \frac{1}{2}n$, where m, n are odd integers. Computing modulo 4, we find

$$m^2 - n^2d \equiv (\pm 1)^2 - (\pm 1)^2 \cdot 1 \equiv 0 \text{ (modulo 4)}.$$

Hence $a^2 - b^2d = \frac{1}{4}(m^2 - n^2d) \in \mathbb{Z}$, as required.

Conversely, suppose that α is an algebraic integer. Then $2a \in \mathbb{Z}$ by Corollary (6.13). There are two cases: either $a \in \mathbb{Z}$ or $a \in \mathbb{Z} + \frac{1}{2}$.

Case 1: $a \in \mathbb{Z}$. It follows that $b^2d \in \mathbb{Z}$ too. Now if we write $b = m/n$, where m, n are relatively prime integers and $n > 0$, then $b^2d = m^2d/n^2$. Since d is square free, it can't cancel a square in the denominator. So $n = 1$. If a is an integer, b must be an integer too.

Case 2: $a \in \mathbb{Z} + \frac{1}{2}$ is a half integer, say $a = \frac{1}{2}m$ as before. Then $4a^2 \in \mathbb{Z}$, and the condition $a^2 - b^2d \in \mathbb{Z}$ implies that $4b^2d \in \mathbb{Z}$ but $b^2d \notin \mathbb{Z}$. Therefore b is also a half integer, say $b = \frac{1}{2}n$, where n is odd. In order for this pair of values for a, b to satisfy $a^2 - b^2d \in \mathbb{Z}$, we must have $m^2 - n^2d \equiv 0$ (modulo 4). Computing modulo 4, we find that $d \equiv 1$ (modulo 4). $\square$

A convenient way to write all the integers in the case $d \equiv 1$ (modulo 4) is to introduce the algebraic integer

(6.15) $$\eta = \tfrac{1}{2}(1 + \delta),$$

which is a root of the monic integer polynomial

(6.16) $$x^2 - x + \tfrac{1}{4}(1 - d).$$

(6.17) **Proposition.** Assume that $d \equiv 1$ (modulo 4). Then the algebraic integers in $F = \mathbb{Q}[\sqrt{d}]$ are $a + b\eta$, where $a, b \in \mathbb{Z}$. $\square$

It is easy to show by explicit calculation that the integers in F form a ring R in each case, called the *ring of integers in F*. Computation in R can be carried out by high school algebra.

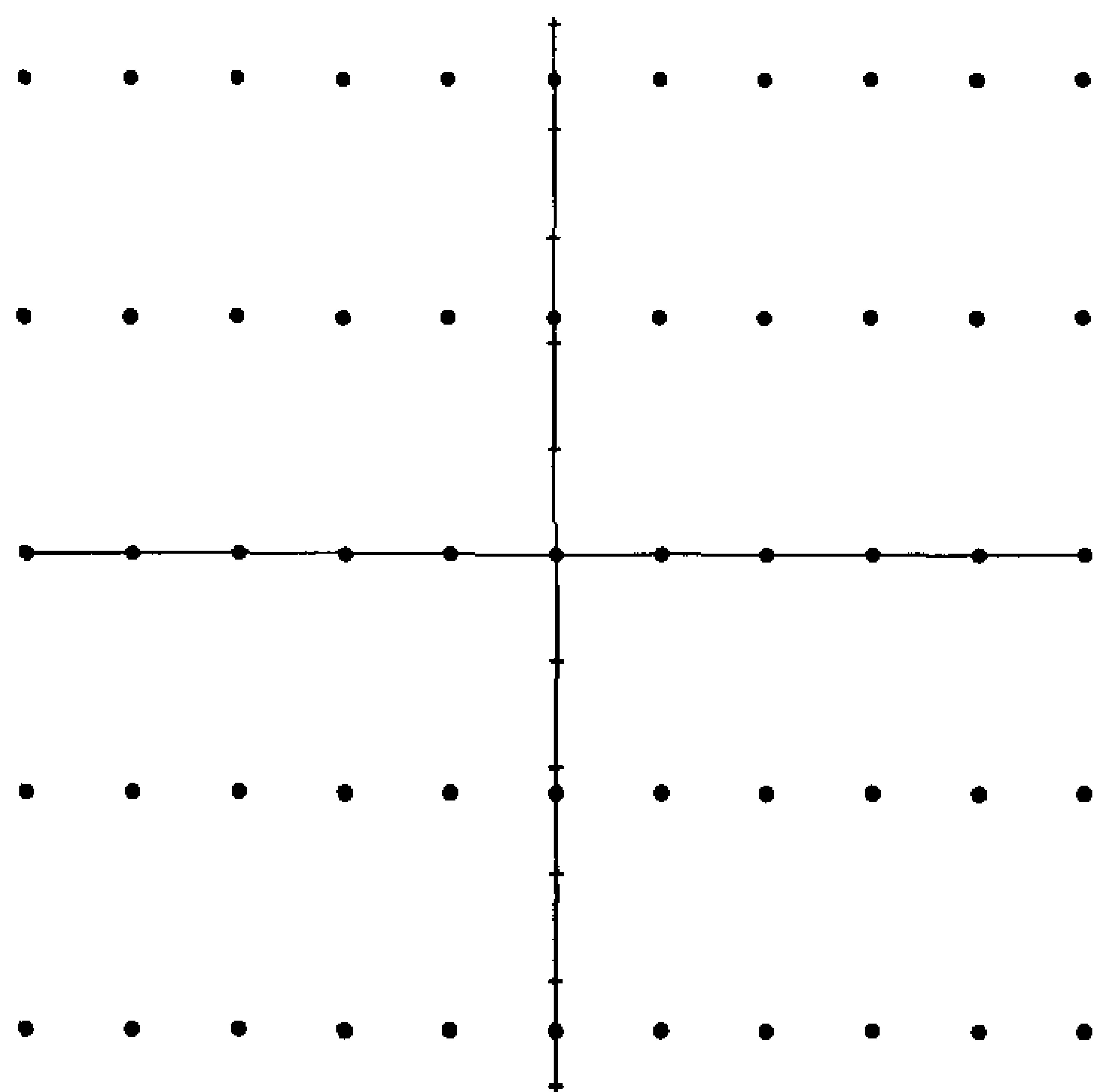
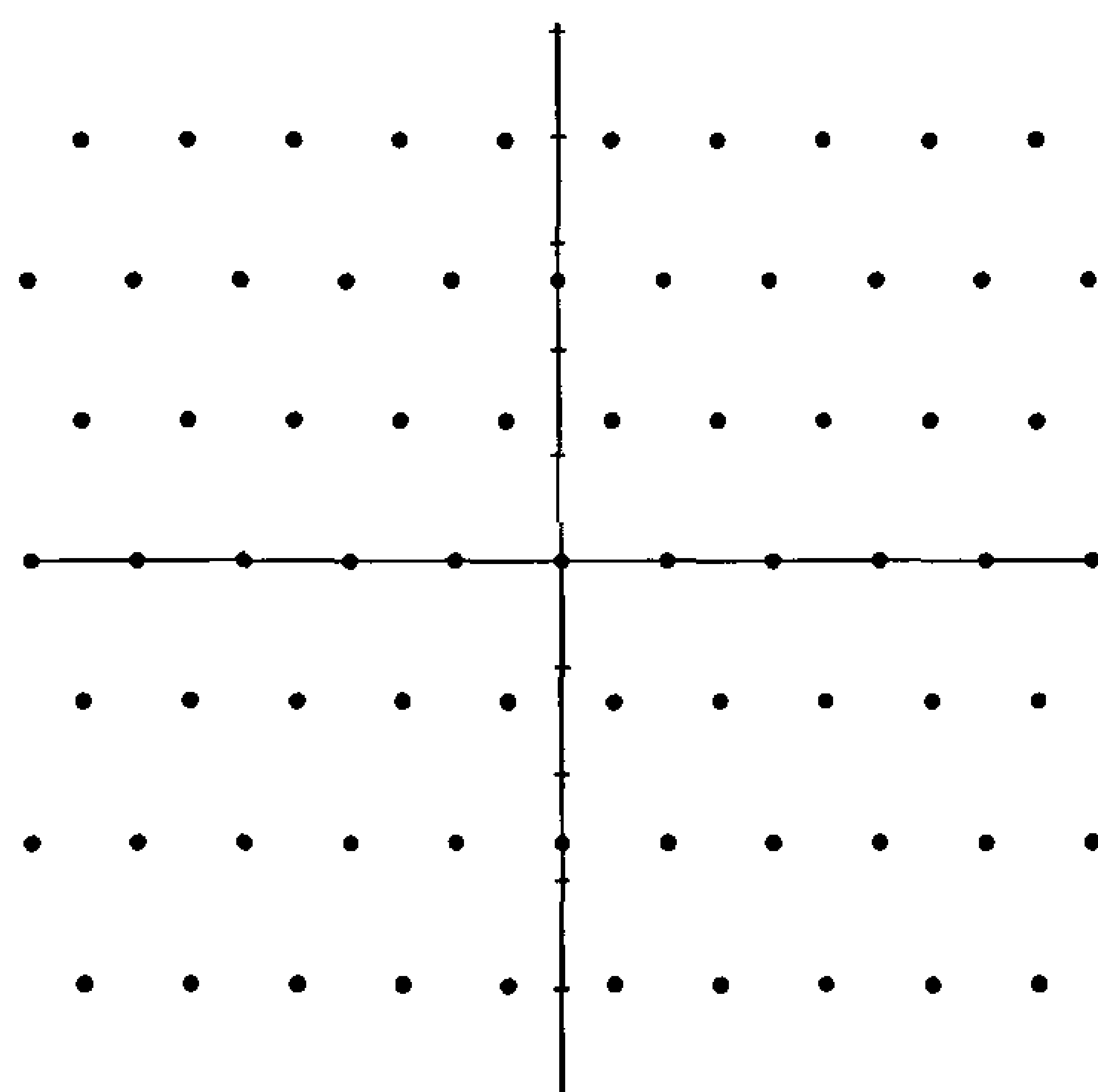
The *discriminant* of F is defined to be the discriminant of the polynomial $x^2 - d$ in the case $R = \mathbb{Z}[\delta]$ and the discriminant of the polynomial $x^2 - x + \frac{1}{4}(1 - d)$ if $R = \mathbb{Z}[\eta]$. This discriminant will be denoted by D. Thus

(6.18) $$D = \begin{cases} 4d & \text{if } d \equiv 2, 3 \\ d & \text{if } d \equiv 1 \end{cases} \quad \text{(modulo 4)}.$$

Since D can be computed in terms of d, it isn't very important to introduce a separate notation for it. However, some formulas become independent of the congruence class when they are expressed in terms of D rather than d.

The imaginary quadratic case $d < 0$ is slightly easier to treat than the real one, so we will concentrate on it in the next sections. In the imaginary case, the ring R

forms a lattice in the complex plane which is rectangular if $d \equiv 2, 3$ (modulo 4), and "isosceles triangular" if $d \equiv 1$ (modulo 4). When $d = -1, R$ is the ring of Gauss integers, and the lattice is square. When $d = -3$, the lattice is equilateral triangular. Two other examples are depicted below.

$d = -5$ $\qquad$ $d = -7$

(6.19) **Figure.** Integers in some imaginary quadratic fields.

The property of being a lattice is very special to rings such as those we are considering here, and we will use geometry to analyze them. Thinking of R as a lattice is also useful for intuition.

It will be helpful to carry along a specific example as we go. We will use the case $d = -5$ for this purpose. Since $-5 \equiv 3$ (modulo 4), the ring of integers forms a rectangular lattice, and $R = \mathbb{Z}[\delta]$, where $\delta = \sqrt{-5}$.

7. FACTORIZATION IN IMAGINARY QUADRATIC FIELDS

Let R be the ring of integers of an imaginary quadratic number field $F = \mathbb{Q}[\delta]$. If $\alpha = a + b\delta$ is in R, so is its complex conjugate $\overline{\alpha} = a - b\delta$. We call the *norm* of α the integer

$$N(\alpha) = \alpha\overline{\alpha}. \tag{7.1}$$

It is also equal to $a^2 - b^2d$ and to $|\alpha|^2$, and it is the constant term of the irreducible polynomial for α over $\mathbb{Q}$. Thus $N(\alpha)$ is a positive integer unless $\alpha = 0$. Note that

$$N(\beta\gamma) = N(\beta)N(\gamma). \tag{7.2}$$

This formula gives us some control of possible factors of an element α of R. Say that $\alpha = \beta\gamma$. Then both terms on the right side of (7.2) are positive integers. So to check for factors of α, it is enough to look at elements β whose norm divides $N(\alpha)$; this is not too big a job if a and b are reasonably small.

In particular, let us ask for *units* of R:

(7.3) **Proposition.**

(a) An element α of R is a unit if and only if $N(\alpha) = 1$.
(b) The units of R are $\{\pm 1\}$ unless $d = -1$ or -3. If $d = -1$, so that R is the ring of Gauss integers, the units are $\{\pm 1, \pm i\}$, and if $d = -3$ they are the powers of the 6th root of unity $\frac{1}{2}(1 + \sqrt{-3})$.

Proof. If α is a unit, then $N(\alpha)N(\alpha^{-1}) = N(1) = 1$. Since $N(\alpha)$ and $N(\alpha^{-1})$ are positive integers, they are both equal to 1. Conversely, if $N(\alpha) = \alpha\overline{\alpha} = 1$, then $\overline{\alpha} = \alpha^{-1}$. So $\alpha^{-1} \in R$, and α is a unit. Thus α is a unit if and only if it lies on the unit circle in the complex plane. The second assertion follows from the configuration of the lattice R [see Figure (6.19)]. □

Next we investigate factorization of an element $\alpha \in R$ into irreducible factors.

(7.4) **Proposition.** Existence of factorizations is true in R.

Proof. If $\alpha = \beta\gamma$ is a proper factorization in R, then β, γ aren't units. So by Proposition (7.3), $N(\alpha) = N(\beta)N(\gamma)$ is a proper factorization in the ring of integers. The existence of factorizations in R now follows from the existence of factorizations in $\mathbb{Z}$. □

However, factorization into irreducible elements will not be unique in most cases. We gave a simple example with $d = -5$ in Section 2:

(7.5)
$$6 = 2 \cdot 3 = (1 + \delta)(1 - \delta),$$

where $\delta = \sqrt{-5}$. For example, to show that $1 + \delta$ is irreducible, we note that its norm is $(1 + \delta)(1 - \delta) = 6$. A proper factor must have norm 2 or 3, that is, absolute value $\sqrt{2}$ or $\sqrt{3}$. There are no such points in the lattice R.

The same method provides examples for other values of d:

(7.6) **Proposition.** The only ring R with $d \equiv 3$ (modulo 4) which is a unique factorization domain is the ring of Gauss integers.

Proof. Assume that $d \equiv 3$ (modulo 4), but that $d \neq -1$. Then

$$1 - d = 2\left(\frac{1 - d}{2}\right) \quad \text{and} \quad 1 - d = (1 + \delta)(1 - \delta).$$

There are two factorizations of $1 - d$ in R. The element 2 is irreducible because $N(2) = 4$ is the smallest value >1 taken on by $N(\alpha)$. [The only points of R inside

the circle of radius 2 about the origin are $0, 1, -1$, when $d = -5, -13, -17, \ldots$. See Figure (6.19).] So if there were a common refinement of the above factorizations, 2 would divide either $1 + \delta$ or $1 - \delta$ in R, which it does not: $\frac{1}{2} \pm \frac{1}{2}\delta$ is not in R when $d \equiv 3$ (modulo 4). □

Notice that this reasoning breaks down if $d \equiv 1$ (modulo 4). In that case, 2 does divide $1 + \delta$, because $\frac{1}{2} + \frac{1}{2}\delta \in R$. In fact, there are more cases of unique factorization when $d \equiv 1$ (modulo 4). The following theorem is very deep, and we will not prove it:

(7.7) **Theorem.** Let R be the ring of integers in the imaginary quadratic field $\mathbb{Q}(\sqrt{d})$. Then R is a unique factorization domain if and only if d is one of the integers $-1, -2, -3, -7, -11, -19, -43, -67, -163$.

Gauss proved for these values of d that R is a unique factorization domain. We will learn how to do this. He also conjectured that there were no others. This much more difficult part of the theorem was finally proved by Baker and Stark in 1966, after the problem had been worked on for more than 150 years.

Ideals were introduced to rescue the uniqueness of factorization. As we know (2.12), R must contain some nonprincipal ideals unless it is a unique factorization domain. We will see in the next section how these nonprincipal ideals serve as substitutes for elements.

Note that every nonzero ideal A is a *sublattice* of R: It is a subgroup under addition, and it is discrete because R is discrete. Moreover, if α is a nonzero element of A, then $\alpha\delta$ is in A too, and $\alpha, \alpha\delta$ are linearly independent over $\mathbb{R}$. However, not every sublattice is an ideal.

(7.8) **Proposition.** If $d \equiv 2$ or 3 (modulo 4), the nonzero ideals of R are the sublattices which are closed under multiplication by δ. If $d \equiv 1$ (modulo 4), they are the sublattices which are closed under multiplication by $\eta = \frac{1}{2}(1 + \delta)$.

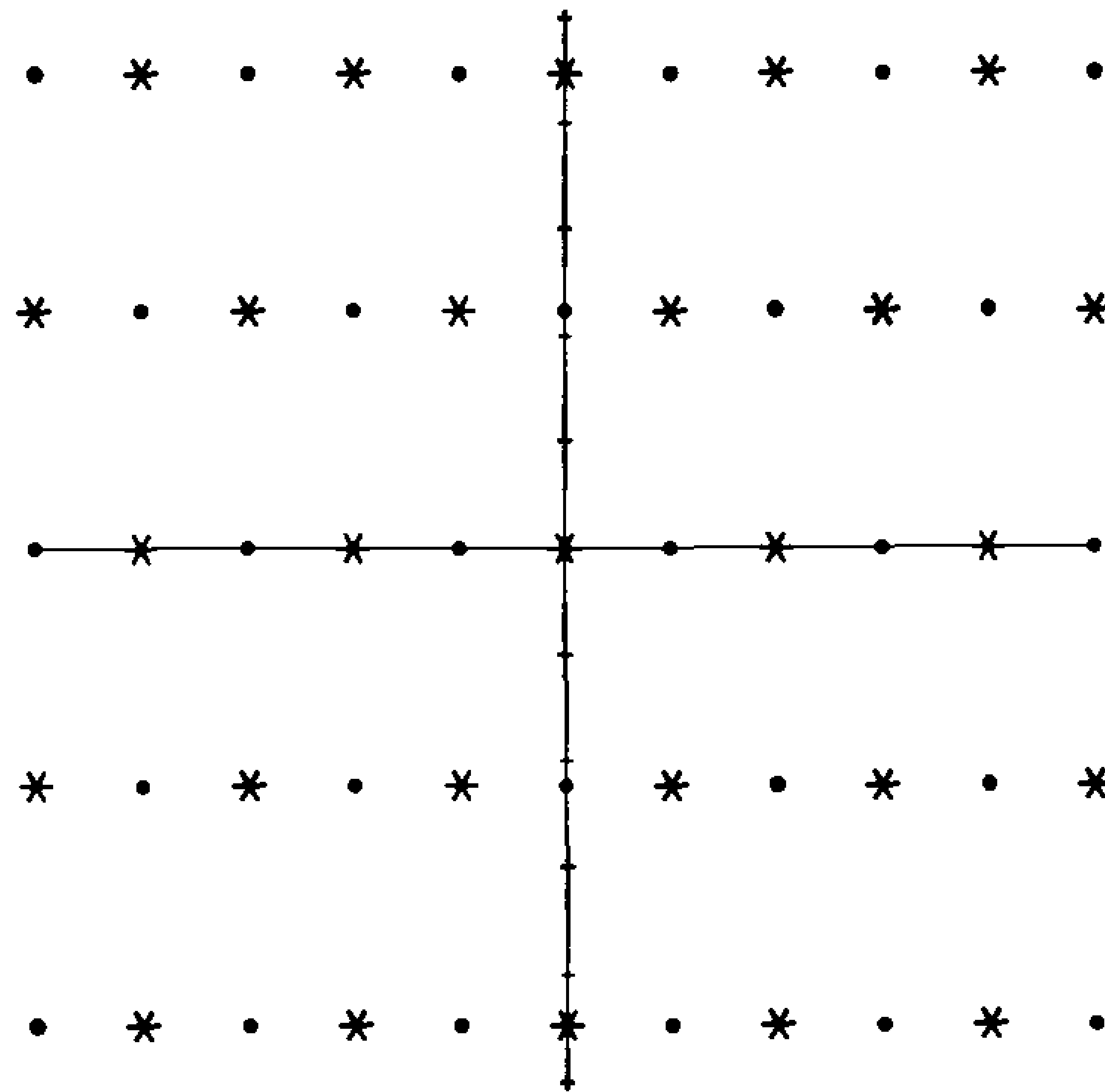
Proof. To be an ideal, a subset A must be closed under addition and under multiplication by elements of R. Any lattice is closed under addition and under multiplication by integers. So if it is also closed under multiplication by δ, then it is also closed under multiplication by an element of the form $a + b\delta$, with $a, b \in \mathbb{Z}$. This includes all elements of R if $d \equiv 2, 3$ (modulo 4). The proof in the case that $d \equiv 1$ (modulo 4) is similar. □

In order to get a feeling for the possibilities, we will describe the ideals of the ring $R = \mathbb{Z}[\sqrt{-5}]$ before going on. The most interesting ideals are those which are not principal.

(7.9) **Theorem.** Let $R = Z[\delta]$, where $\delta = \sqrt{-5}$, and let A be a nonzero ideal of R. Let α be a nonzero element of A of minimal absolute value $|\alpha|$. There are two cases:

Case 1: A is the principal ideal (α), which has the lattice basis $(\alpha, \alpha\delta)$.
Case 2: A has the lattice basis $(\alpha, \frac{1}{2}(\alpha + \alpha\delta))$, and is not a principal ideal.

The second case can occur only if $\frac{1}{2}(\alpha + \alpha\delta)$ is an element of R. The ideal $A = (2, 1 + \delta)$, which is depicted below, is an example.



(7.10) **Figure.** The ideal $(2, 1 + \delta)$ in the ring $\mathbb{Z}[\delta]$, $\delta = \sqrt{-5}$.

The statement of Proposition (7.9) has a geometric interpretation. Notice that the lattice basis $(\alpha, \alpha\delta)$ of the principal ideal (α) is obtained from the lattice basis $(1, \delta)$ of R by multiplication by α. If we write $\alpha = re^{i\theta}$, then the effect of multiplication by α is to rotate the complex plane through the angle θ and then stretch by the factor r. So (α) and R are similar geometric figures, as we noted in Section 2. Similarly, the basis $(\alpha, \frac{1}{2}(\alpha + \alpha\delta))$ is obtained by multiplication by $\frac{1}{2}\alpha$ from the basis $(2, 1 + \delta)$. So the ideals listed in Case 2 are geometric figures similar to the one depicted in Figure (7.10). The similarity classes of ideals are called the *ideal classes*, and their number is called the *class number* of R. Thus Proposition (7.9) implies that the class number of $\mathbb{Z}[\sqrt{-5}]$ is 2. We will discuss ideal classes for other quadratic imaginary fields in Section 10.

The proof of Theorem (7.9) is based on the following lemma about lattices in the complex plane:

(7.11) **Lemma.** Let r be the minimum absolute value among nonzero elements of a lattice A, and let γ be an element of A. Let D be the disc of radius $\frac{1}{n}r$ about the point $\frac{1}{n}\gamma$. There is no point of A in the interior of D other than its center $\frac{1}{n}\gamma$.

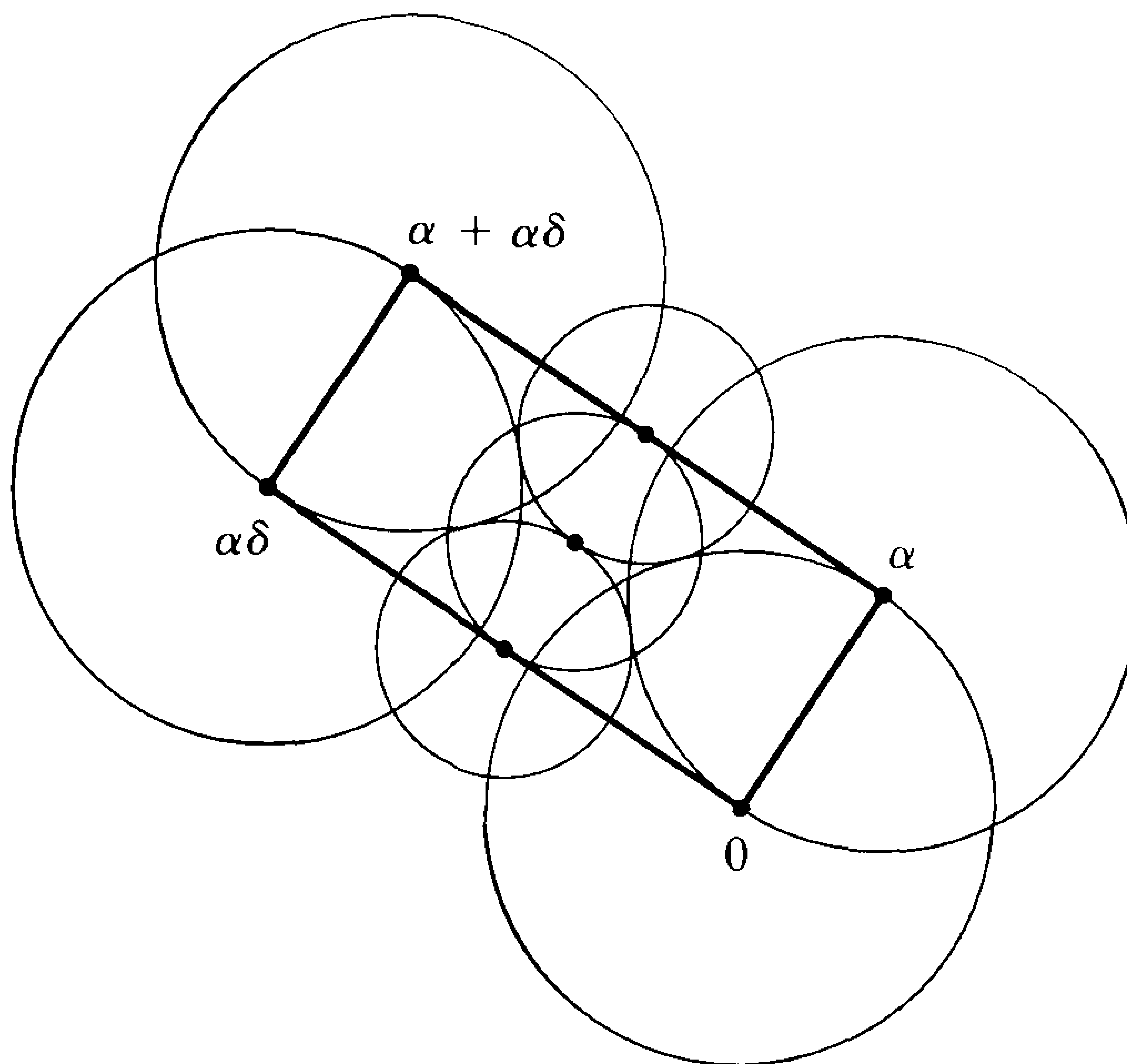
The point $\frac{1}{n}\gamma$ may lie in A or not. This depends on A and on γ.

Proof. Let β be a point in the interior of D. Then by definition of the disc, $|\beta - \frac{1}{n}\gamma| < \frac{1}{n}r$, or equivalently, $|n\beta - \gamma| < r$. If $\beta \in A$, then $n\beta - \gamma \in A$ too.

In this case, $n\beta - \gamma$ is an element of A of absolute value less than r, which implies that $n\beta - \gamma = 0$, hence that $\beta = \frac{1}{n}\gamma$. □

Proof of Theorem (7.9). Let α be the chosen element of A of minimal absolute value r. The principal ideal $(\alpha) = R\alpha$ consists of the complex numbers $(a + b\delta)\alpha$, with $a, b \in \mathbb{Z}$. So it has the lattice basis $(\alpha, \alpha\delta)$ as is asserted in the proposition. Since A contains α, it contains the principal ideal (α) too, and if $A = (\alpha)$ we are in Case 1.

Suppose that $A > (\alpha)$, and let β be an element of A which is not in (α). We may choose β to lie in the rectangle whose four vertices are $0, \alpha, \alpha\delta, a + a\delta$ [see Chapter 5 (4.14)]. Figure (7.13) shows a disc of radius r about the four vertices of this rectangle, and a disc of radius $\frac{1}{2}r$ about the three half lattice points $\frac{1}{2}\alpha\delta, \frac{1}{2}(\alpha + \alpha\delta)$, and $\alpha + \frac{1}{2}\alpha\delta$. Notice that the interiors of these discs cover the rectangle. According to Lemma (7.11), the only points of the interiors which can lie in A are the centers of the discs. Since β is not in (α), it is not a vertex of the rectangle. So β must be one of the half lattice points $\frac{1}{2}\alpha\delta, \frac{1}{2}(\alpha + \alpha\delta)$, or $\alpha + \frac{1}{2}\alpha\delta$.



(7.13) **Figure.**

This exhausts the information which we can get from the fact that A is a lattice. We now use the fact that A is an ideal to rule out the two points $\frac{1}{2}\alpha\delta$ and $\alpha + \frac{1}{2}\alpha\delta$. Suppose that $\frac{1}{2}\alpha\delta \in A$. Multiplying by δ, we find that $\frac{1}{2}\alpha\delta^2 = -\frac{5}{2}\alpha \in A$ too and since $\alpha \in A$ that $\frac{1}{2}\alpha \in A$. This contradicts our choice of α. Next, we note that if $\alpha + \frac{1}{2}\alpha\delta$ were in A, then $\frac{1}{2}\alpha\delta$ would be in A too, which has been ruled out. The remaining possibility is that $\beta = \frac{1}{2}(\alpha + \alpha\delta)$. If so, we are in Case 2. □

8. IDEAL FACTORIZATION

Let R be the ring of integers in an imaginary quadratic field. In order to avoid confusion, we will denote ordinary integers by latin letters $a, b, \ldots$, elements of R by greek letters $\alpha, \beta, \ldots$, and ideals by capital letters $A, B, \ldots$. We will consider only *nonzero* ideals of R.

The notation $A = (\alpha, \beta, \ldots, \gamma)$ stands for the ideal generated by the elements $\alpha, \beta, \ldots, \gamma$. Since an ideal is a plane lattice, it has a lattice basis consisting of two elements. Any lattice basis generates the ideal, but we must distinguish between the notions of a lattice basis and a generating set. We also need to remember the dictionary (2.2) which relates elements to the principal ideals they generate.

Dedekind extended the notion of divisibility to ideals using the following definition of ideal multiplication: Let A and B be ideals in a ring R. We would like to define the product ideal AB to be the set of all products $\alpha\beta$, where $\alpha \in A$ and $\beta \in B$. Unfortunately, this set of products is usually not an ideal: It will not be closed under sums. To get an ideal, we must put into AB all *finite sums of products*

$$(8.1) \qquad \sum_i \alpha_i \beta_i, \quad \text{where } \alpha_i \in A \text{ and } \beta_i \in B.$$

The set of such sums is the smallest ideal of R which contains all products $\alpha\beta$, and we denote this *product ideal* by AB. (This use of the product notation is different from its use in group theory [Chapter 2 (8.5)].) The definition of multiplication of ideals is not as simple as we might hope, but it works reasonably well.

Notice that multiplication of ideals is commutative and associative, and that R is a unit element. This is why $R = (1)$ is often called the unit ideal:

$$(8.2) \qquad AR = RA = A, \quad AB = BA, \quad A(BC) = (AB)C.$$

(8.3) **Proposition.**

(a) The product of principal ideals is principal: If $A = (\alpha)$ and $B = (\beta)$, then $AB = (\alpha\beta)$.

(b) Assume that $A = (\alpha)$ is principal, but let B be arbitrary. Then

$$AB = \alpha B = \{\alpha\beta \mid \beta \in B\}.$$

(c) Let $\alpha_1, \ldots, \alpha_m$ and $\beta_1, \ldots, \beta_n$ be generators for the ideals A and B respectively. Then AB is generated as an ideal by the mn products $\alpha_i\beta_j$.

We leave this proof as an exercise. □

In analogy with divisibility of elements of a ring, we say that an ideal A *divides* another ideal B if there is an ideal C such that $B = AC$.

To see how multiplication of ideals can be used, let us go back to the example $d = -5$, in which $2 \cdot 3 = (1 + \delta)(1 - \delta)$. For uniqueness of factorization to hold in the ring $R = \mathbb{Z}[\delta]$, there would have to be an element $\rho \in R$ dividing both 2 and

$1 + \delta$. This is the same as saying that 2 and $1 + \delta$ should be in the principal ideal (ρ). There is no such element. However, there is an *ideal*, not a principal ideal, which contains 2 and $1 + \delta$, namely the ideal generated by these two elements. This ideal $A = (2, 1 + \delta)$ is depicted in Figure (7.10). We can make three other ideals using the factors of 6:

$$\overline{A} = (2, 1 - \delta), \quad B = (3, 1 + \delta), \quad \overline{B} = (3, 1 - \delta).$$

The first of these ideals is denoted by $\overline{A}$ because it is the complex conjugate of the ideal A:

$$\overline{A} = \{\overline{\alpha} \mid \alpha \in A\}. \tag{8.4}$$

As a lattice, $\overline{A}$ is obtained by reflecting the lattice A about the real axis. That the complex conjugate of any ideal is an ideal is easily seen. Actually, it happens that our ideal A is equal to its complex conjugate $\overline{A}$, because $1 - \delta = 2 - (1 + \delta) \in A$. This is an accidental symmetry of the lattice A: The ideals B and $\overline{B}$ are not the same.

Now let us compute the products of these ideals. According to Proposition (8.3c), the ideal $A\overline{A}$ is generated by the four products of the generators $(2, 1 + \delta)$ and $(2, 1 - \delta)$ of A and $\overline{A}$:

$$A\overline{A} = (4, 2 + 2\delta, 2 - 2\delta, 6).$$

Each of these four generators is divisible by 2, so $A\overline{A} \subset (2)$. On the other hand, $2 = 6 - 4$ is in $A\overline{A}$. Therefore $(2) \subset A\overline{A}$, so

$$A\overline{A} = (2)\,!$$

[The notation (2) is ambiguous, because it can denote both $2R$ and $2\mathbb{Z}$. It stands for $2R$ here.] Next, the product AB is generated by the four products:

$$AB = (6, 2 + 2\delta, 3 + 3\delta, -4 + 2\delta).$$

Each of these four elements is divisible by $1 + \delta$. Since $1 + \delta$ is in AB, we find that $AB = (1 + \delta)$. Similarly, $\overline{AB} = (1 - \delta)$ and $B\overline{B} = (3)$.

It follows that the principal ideal (6) is the product of the four ideals:

$$(6) = (2)(3) = (A\overline{A})(B\overline{B}) = (AB)(\overline{AB}) = (1 + \delta)(1 - \delta). \tag{8.5}$$

Isn't this beautiful? The ideal factorization $(6) = A\overline{A}B\overline{B}$ has provided a common refinement of the two factorizations (2.7).

The rest of this section is devoted to proving unique factorization of ideals in the rings of integers of an imaginary quadratic number field. We will follow the discussion of factorization of elements as closely as possible.

The first thing to do is to find an analogue for ideals of the notion of a prime element.

(8.6) **Proposition.** Let P be an ideal of a ring R which is not the unit ideal. The following conditions are equivalent:

(i) If α, β are elements of R such that $\alpha\beta \in P$, then $\alpha \in P$ or $\beta \in P$.

(ii) If A, B are ideals of R such that $AB \subset P$, then $A \subset P$ or $B \subset P$.
(iii) The quotient ring R/P is an integral domain.

An ideal which satisfies one of these conditions is called a *prime ideal.*

For example, every maximal ideal is prime, because if M is maximal, then R/M is a field, and a field is an integral domain. The zero ideal of a ring R is prime if and only if R is an integral domain.

Proof of the Proposition: The conditions for $\overline{R} = R/P$ to be an integral domain are that $\overline{R} \neq 0$ and that $\overline{\alpha\beta} = 0$ implies $\overline{\alpha} = 0$ or $\overline{\beta} = 0$. These conditions translate back to $P \neq R$ and if $\alpha\beta \in P$ then $\alpha \in P$ or $\beta \in P$. Thus (i) and (iii) are equivalent. The fact that (ii) implies (i) is seen by taking $A = (\alpha)$ and $B = (\beta)$. The only surprising implication is that (i) implies (ii). Assume that (i) holds, and let A, B be ideals such that $AB \subset P$. If A is not contained in P, there is some element $\alpha \in A$ which is not in P. If β is an element of B, then $\alpha\beta \in AB$; hence $\alpha\beta \in P$. By part (i), $\beta \in P$. Since this is true for all of its elements, $B \subset P$ as required. □

We now go back to imaginary quadratic number fields.

(8.7) **Lemma.** Let $A \subset B$ be lattices in $\mathbb{R}^2$. There are only finitely many lattices L between A and B, that is, such that $A \subset L \subset B$.

Proof. Let (α_1, α_2) be a lattice basis for A, and let P be the parallelogram with vertices $0, \alpha_1, \alpha_2, \alpha_1 + \alpha_2$. There are finitely many elements of B contained in P [Chapter 5 (4.12)], so if L is a lattice between A and B, there are finitely many possibilities for the set $L \cap P$. Call this set S. The proof will be completed by showing that S and A determine the lattice L. To show this, let γ be an element of L. Then there is an element of $\alpha \in A$ such that $\gamma - \alpha$ is in P, hence in S. [See the proof of (4.14) in Chapter 5]. Symbolically, we have $L = S + A$. This describes L in terms of S and A, as required. □

(8.8) **Proposition.** Let R be the ring of integers in an imaginary quadratic number field.

(a) Let B be a nonzero ideal of R. There are finitely many ideals between B and R.
(b) Every proper ideal of R is contained in a maximal ideal.
(c) The nonzero prime ideals of R are the maximal ideals.

Proof.

(a) This follows from lemma (8.7).

(b) Let B be a proper ideal. Then B is contained in only finitely many ideals. We can search through them to find a maximal ideal.

(c) We have already remarked that maximal ideals are prime. Conversely, let P be a nonzero prime ideal. Then P has finite index in R. So R/P is a finite integral do-

main, and hence it is a field [Chapter 10 (6.4)]. This shows that P is a maximal ideal. □

(8.9) **Theorem.** Let R be the ring of integers in an imaginary quadratic field F. Every nonzero ideal of R which is not the whole ring is a product of prime ideals. This factorization is unique, up to order of the factors.

This remarkable theorem can be extended to other rings of algebraic integers, but it is a very special property of such rings. Most rings do not admit unique factorization of ideals. Several things may fail, and we want to take particular note of one of them. We know that a principal ideal (α) contains another principal ideal (β) if and only if α divides β in the ring. So the definition of a prime element π can be restated as follows: If $(\pi) \supset (\alpha\beta)$, then $(\pi) \supset (\alpha)$ or $(\pi) \supset (\beta)$. The second of the equivalent definitions (8.6) of a prime ideal is the analogous statement for ideals: If $P \supset AB$, then $P \supset A$ or $P \supset B$. So if inclusion of ideals were equivalent with divisibility, the proof of uniqueness of factorizations would carry over to ideals. Unfortunately the cumbersome definition of product ideal causes trouble. In most rings, the inclusion $A \supset B$ does not imply that A divides B. This weakens the analogy between prime ideal and prime element. It will be important to establish the equivalence of inclusion and divisibility in the particular rings we are studying. This is done below, in Proposition (8.11).

We now proceed with the proof of Theorem (8.9). For the rest of this section, R will denote the ring of integers in an imaginary quadratic number field. The proof is based on the following lemma:

(8.10) **Main Lemma.** Let R be the ring of integers in an imaginary quadratic number field. The product of a nonzero ideal and its conjugate is a principal ideal of R generated by an ordinary integer:

$$A\overline{A} = (n), \quad \text{for some } n \in \mathbb{Z}.$$

The most important point here is that for every ideal A there is some ideal B such that AB is principal. That $\overline{A}$ does the job and that the product ideal is generated by an ordinary integer are less important points.

We will prove the lemma at the end of the section. Let us assume it for now and derive some consequences for multiplication of ideals. Because these consequences depend on the Main Lemma, they are not true for general rings.

(8.11) **Proposition.** Let R be the ring of integers in an imaginary quadratic number field.

(a) *Cancellation Law:* Let A, B, C be nonzero ideals of R. If $AB \supset AC$ then $B \supset C$. If $AB = AC$, then $B = C$.

(b) If A and B are nonzero ideals of R, then $A \supset B$ if and only if A divides B, that is, if and only if $B = AC$ for some ideal C.

(c) Let P be a nonzero prime ideal of R. If P divides a product AB of ideals, then P divides one of the factors A or B.

Proof. (a) Assume that $AB \supset AC$. If $A = (\alpha)$ is principal, then $AB = \alpha B$ and $AC = \alpha C$ (8.3). Viewing these sets as subsets of the complex numbers, we multiply the relation $\alpha B \supset \alpha C$ on the left by α^{-1} to conclude that $B \supset C$. So the assertion is true when A is principal. In general, if $AB \supset AC$, then multiply both sides by $\bar{A}$ and apply the Main Lemma: $nB = \bar{A}AB \supset \bar{A}AC = nC$, and apply what has been shown. The case that $AB = AC$ is the same.

(b) The implication which is not clear is that if A contains B then A divides B. We will first check this when $A = (\alpha)$ is principal. In this case, to say that $(\alpha) \supset B$ means that α divides every element β of B. Let $C = \alpha^{-1}B$ be the set of quotients, that is, the set of elements $\alpha^{-1}\beta$, with $\beta \in B$. You can check that C is an ideal and that $\alpha C = B$. Hence $B = AC$ in this case. Now let A be arbitrary, and assume that $A \supset B$. Then $(n) = \bar{A}A \supset \bar{A}B$. By what has already been shown, there is an ideal C such that $nC = \bar{A}B$, or $\bar{A}AC = \bar{A}B$. By the Cancellation Law, $AC = B$.

(c) To prove part (c) of the proposition, we apply part (b) to translate divisibility into inclusion. Then (c) follows from the definition of prime ideal. □

Proof of Theorem (8.9). There are two things to prove. First we must show that every proper, nonzero ideal A is a product of prime ideals. If A is not itself prime, then it is not maximal, so we can find a proper ideal A_1 strictly larger than A. Then A_1 divides A (8.11b), so we can write $A = A_1 B_1$. It follows that $A \subset B_1$. Moreover, if we had $A = B_1$, the Cancellation Law would imply $R = A_1$, contradicting the fact that A_1 is a proper ideal. Thus $A < B_1$. Similarly, $A < A_1$. Since there are only finitely many ideals between A and R, this process of factoring an ideal terminates. When it does, all factors will be maximal, and hence prime. So every proper ideal A can be factored into primes.

Now to prove uniqueness, we apply the property (8.11c) of prime ideals: If $P_1 \cdots P_r = Q_1 \cdots Q_s$, with P_i, Q_j prime, then P_1 divides $Q_1 \cdots Q_s$, and hence it divides one of the factors, say Q_1. Since Q_1 is maximal, $P_1 = Q_1$. Cancel by (8.11a) and use induction on r. □

(8.12) **Theorem.** The ring of integers R is a unique factorization domain if and only if it is a principal ideal domain. If so, then the factorizations of elements and of ideals correspond naturally.

Proof. We already know that a principal ideal domain has unique factorization (2.12). Conversely, suppose that R is a unique factorization domain, and let P be any nonzero prime ideal of R. Then P contains an irreducible element, say π. For, any nonzero element α of P is a product of irreducible elements, and, by definition of prime ideal, P contains one of its irreducible factors. By (2.8), an irreducible element π is prime, that is, (π) is a prime ideal. By (8.6), (π) is maximal. Since

$(\pi) \subset P$, it follows that $(\pi) = P$, hence that P is principal. By Theorem (8.9), every nonzero ideal A is a product of primes; hence it is principal (8.3a). Thus R is a principal ideal domain. The last assertion of the theorem is clear from (2.2). □

Proof of the Main Lemma (8.10). We can generate A as a lattice by two elements, say α, β. Then A is certainly generated as an ideal by these same elements, and moreover $\overline{\alpha}, \overline{\beta}$ generate $\overline{A}$. Hence the four products $\alpha\overline{\alpha}$, $\alpha\overline{\beta}$, $\overline{\alpha}\beta$, $\beta\overline{\beta}$ generate the ideal $A\overline{A}$. Consider the three elements $\alpha\overline{\alpha}$, $\beta\overline{\beta}$, and $\alpha\overline{\beta} + \overline{\alpha}\beta$ of $A\overline{A}$. They are all equal to their conjugates and hence are rational numbers. Since they are algebraic integers, they are ordinary integers. Let n be their greatest common divisor in $\mathbb{Z}$. Then n is a linear combination of $\alpha\overline{\alpha}$, $\beta\overline{\beta}$, $\alpha\overline{\beta} + \overline{\alpha}\beta$ with integer coefficients. Hence n is in the product ideal $A\overline{A}$. Therefore $A\overline{A} \supset (n)$. If we show that n divides each of the four generators of the ideal $A\overline{A}$ in R, then it will follow that $(n) \supset A\overline{A}$, hence that $(n) = A\overline{A}$, as was to be shown.

Now by construction, n divides $\alpha\overline{\alpha}$ and $\beta\overline{\beta}$ in $\mathbb{Z}$, hence in R. So we have to show that n divides $\alpha\overline{\beta}$ and $\overline{\alpha}\beta$ in R. The elements $(\alpha\overline{\beta})/n$ and $(\overline{\alpha}\beta)/n$ are roots of the polynomial $x^2 - rx + s$, where

$$r = \frac{\alpha\overline{\beta} + \overline{\alpha}\beta}{n} \quad \text{and} \quad s = \frac{\alpha\overline{\alpha}}{n}\frac{\beta\overline{\beta}}{n}.$$

By definition of n, these elements r, s are integers, so this is a monic equation in $\mathbb{Z}[x]$. Hence $(\alpha\overline{\beta})/n$ and $(\overline{\alpha}\beta)/n$ are algebraic integers, as required. □

Note. This is the only place where the definition of algebraic integer is used directly. The lemma would be false if we took a smaller ring than R, for example, if we didn't take the elements with half integer coefficients when $d \equiv 1$ (modulo 4).

9. THE RELATION BETWEEN PRIME IDEALS OF R AND PRIME INTEGERS

We saw in Section 5 how the primes in the ring of Gauss integers are related to integer primes. A similar analysis can be made for the ring R of integers in a quadratic number field. The main difference is that R is usually not a principal ideal domain, and therefore we should speak of prime ideals rather than of prime elements. This complicates the analogues of parts (c) and (d) of Theorem (5.1), and we will not consider them here. [However, see (12.10).]

(9.1) **Proposition.** Let P be a nonzero prime ideal of R. There is an integer prime p so that either $P = (p)$ or $P\overline{P} = (p)$. Conversely, let p be a prime integer. There is a prime ideal P of R so that either $P = (p)$ or $P\overline{P} = (p)$.

The proof follows that of parts (a) and (b) of Theorem (5.1) closely. □

The second case of (9.1) is often subdivided into two cases, according to whether or not P and $\overline{P}$ are equal. The following terminology is customary: If (p) is a prime ideal, then we say that p *remains prime* in R. If $P\overline{P} = (p)$, then we say that p *splits* in R, unless $P = \overline{P}$, in which case we say that P *ramifies* in R.

Let us analyze the behavior of primes further. Assume that $d \equiv 2$ or 3 (modulo 4). In this case, $R = \mathbb{Z}[\delta]$ is isomorphic to $\mathbb{Z}[x]/(x^2 - d)$. To ask for prime ideals containing the ideal (p) is equivalent to asking for prime ideals of the ring $R/(p)$ [Chapter 10 (4.3)]. Note that

$$R/(p) \approx \mathbb{Z}[x]/(x^2 - d, p). \tag{9.2}$$

Interchanging the order of the two relations $x^2 - d = 0$ and $p = 0$ as in the proof of Theorem (5.1), we find the first part of the proposition below. The second part is obtained in the same way, using the polynomial (6.16).

(9.3) **Proposition.**

(a) Assume that $d \equiv 2$ or 3 (modulo 4). An integer prime p remains prime in R if and only if the polynomial $x^2 - d$ is irreducible over $\mathbb{F}_p$.

(b) Assume that $d \equiv 1$ (modulo 4). Then p remains prime if and only if the polynomial $x^2 - x + \frac{1}{4}(1 - d)$ is irreducible over $\mathbb{F}_p$. □

10. IDEAL CLASSES IN IMAGINARY QUADRATIC FIELDS

As before, R denotes the ring of integers in an imaginary quadratic number field. In order to analyze the extent to which uniqueness of factorization of elements fails in R, we introduce an equivalence relation on ideals which is compatible with ideal multiplication and such that the principal ideals form one equivalence class. It is reasonably clear which relation to use: We call two ideals A, B *similar* $(A \sim B)$ if there are nonzero elements $\sigma, \tau \in R$ so that

$$\sigma B = \tau A. \tag{10.1}$$

This is an equivalence relation. The equivalence classes for this relation are called *ideal classes*, and the ideal class of A will be denoted by $\langle A \rangle$.

We could also take the element $\lambda = \sigma^{-1}\tau$ of the quadratic number field $F = \mathbb{Q}[\delta]$ and say that A and B are similar if

$$B = \lambda A, \quad \text{for some } \lambda \in \mathbb{Q}[\delta]. \tag{10.2}$$

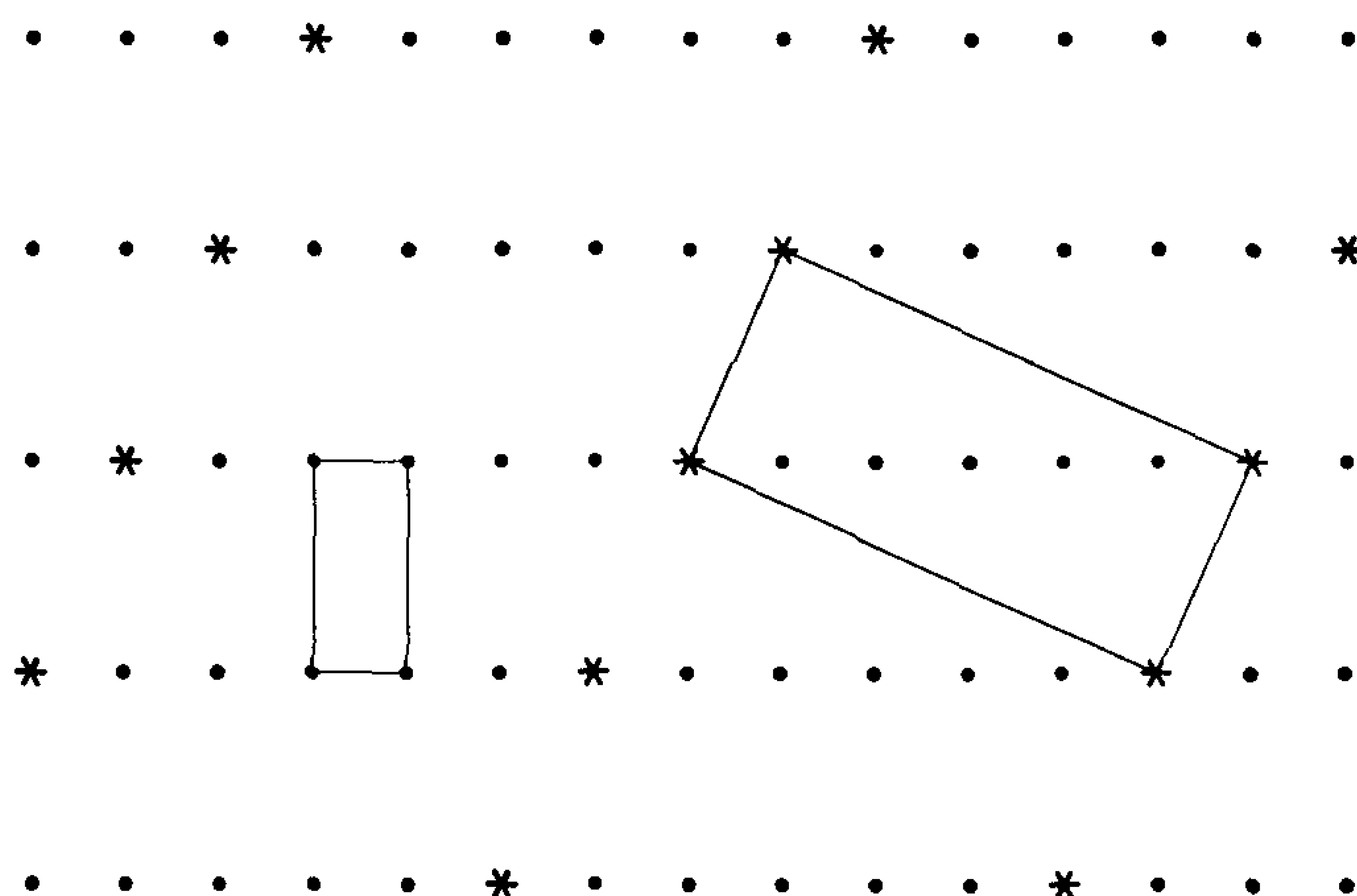
Similarity has a nice geometric interpretation. Two ideals A and B are similar if the lattices in the complex plane which represent them are similar geometric figures, by a similarity which is *orientation-preserving*. To see this, note that a lattice looks the same at all points. So a similarity can be assumed to relate 0 in A to 0 in B. Then it will be described as a rotation followed by a stretching or shrinking,

that is, as multiplication by a complex number λ. Since multiplication by λ carries a nonzero element $\alpha \in A$ to an element $\lambda\alpha = \beta \in B$, $\lambda = \beta\alpha^{-1}$ is automatically in the field F.

An ideal B is similar to the unit ideal R if and only if $B = \lambda R$ for some λ in the field. Then λ is an element of B, hence of R. In this case, B is the principal ideal (λ). So we have the following:

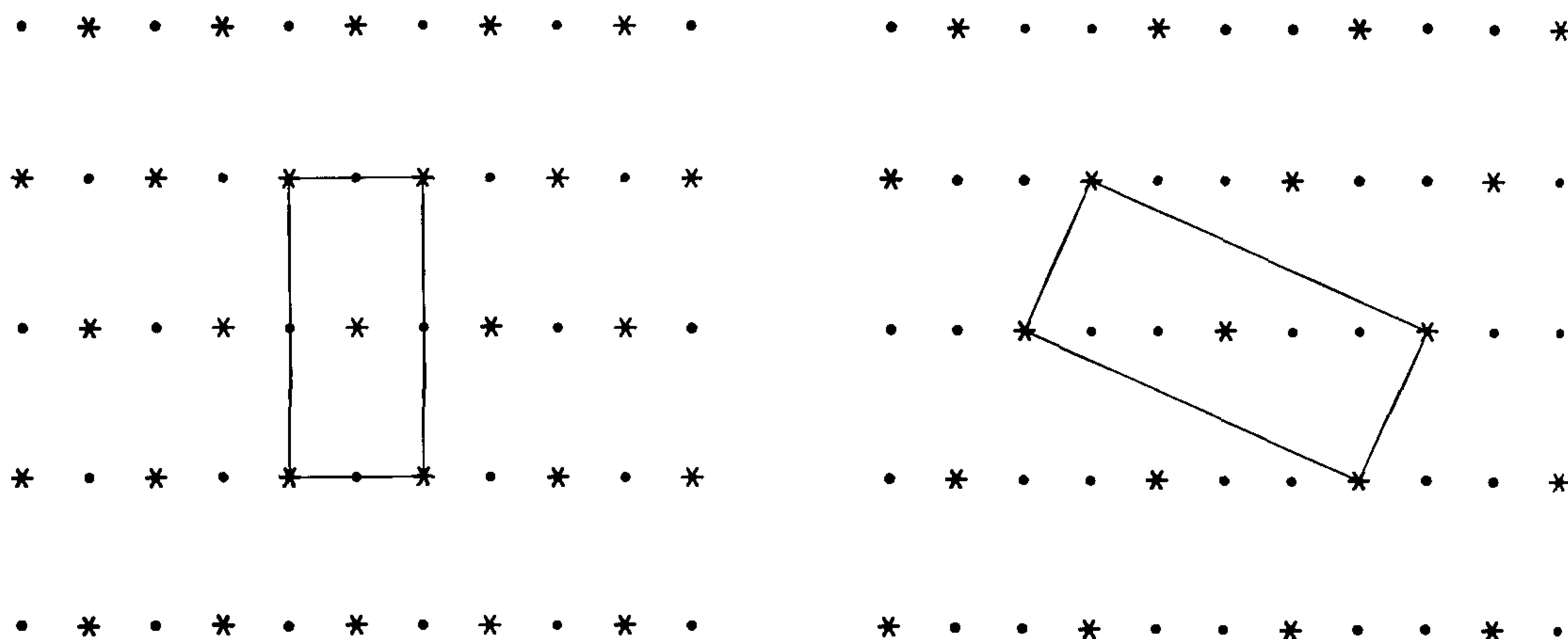
(10.3) **Proposition.** The ideal class $\langle R \rangle$ consists of the principal ideals. □

Figure (10.4) shows the principal ideal $(1 + \delta)$ in the ring $\mathbb{Z}[\delta]$, where $\delta^2 = -5$.



(10.4) **Figure.** The principal ideal $1 + \delta$.

We saw in (7.9) that there are two ideal classes. Each of the ideals $A = (2, 1 + \delta)$ and $B = (3, 1 + \delta)$, for example, represents the class of nonprincipal ideals. In this case $2B = (1 + \delta)A$. These ideals are depicted in Figure (10.5).



(10.5) **Figure.** The ideals $(2,\ 1 + \delta)$ and $(3,\ 1 + \delta)$.

(10.6) **Proposition.** The ideal classes form an abelian group $\mathscr{C}$, with law of composition induced by multiplication of ideals:

$$\langle A\rangle\langle B\rangle = \text{class of } AB = \langle AB\rangle;$$

the class of the principal ideals is the identity: $\langle R\rangle = \langle 1\rangle$.

Proof. If $A \sim A'$ and $B \sim B'$, then $A' = \lambda A$ and $B' = \mu B$ for some $\lambda, \mu \in F = \mathbb{Q}[\delta]$; hence $A'B' = \lambda\mu AB$. This shows that $\langle AB\rangle = \langle A'B'\rangle$, hence that this law of composition is well-defined. Next, the law is commutative and associative because multiplication of ideals is, and the class of R is an identity (8.2). Finally, $A\bar{A} = (n)$ is principal by the Main Lemma (8.10). Since the class of the principal ideal (n) is the identity in $\mathscr{C}$, we have $\langle A\rangle\langle \bar{A}\rangle = \langle R\rangle$, so $\langle \bar{A}\rangle = \langle A\rangle^{-1}$. □

(10.7) **Corollary.** Let R be the ring of integers in an imaginary quadratic number field. The following assertions are equivalent:

(i) R is a principal ideal domain;
(ii) R is a unique factorization domain;
(iii) the ideal class group $\mathscr{C}$ of R is the trivial group.

For to say that $\mathscr{C}$ is trivial is the same as saying that every ideal is similar to the unit ideal, which by Proposition (10.3) means that every ideal is principal. By Theorem (8.12), this occurs if and only if R is a unique factorization domain. □

Because of Corollary (10.7), it is natural to count the ideal classes and to consider this count, called the *class number*, a measure of nonuniqueness of factorization of elements in R. More precise information is given by the structure of $\mathscr{C}$ as a group. As we have seen (7.9), there are two ideal classes in the ring $\mathbb{Z}[\sqrt{-5}]$, so its ideal class group is a cyclic group of order 2 and its class number is 2.

We will now show that the ideal class group $\mathscr{C}$ is always a finite group. The proof is based on a famous lemma of Minkowski about lattice points in convex regions. A bounded subset S of the plane $\mathbb{R}^2$ is called *convex* and *centrally symmetric* if it has these properties:

(10.8) (a) *Convexity:* If $p, q \in S$, then the line segment joining p to q is in S.
(b) *Central symmetry:* If $p \in S$, then $-p \in S$.

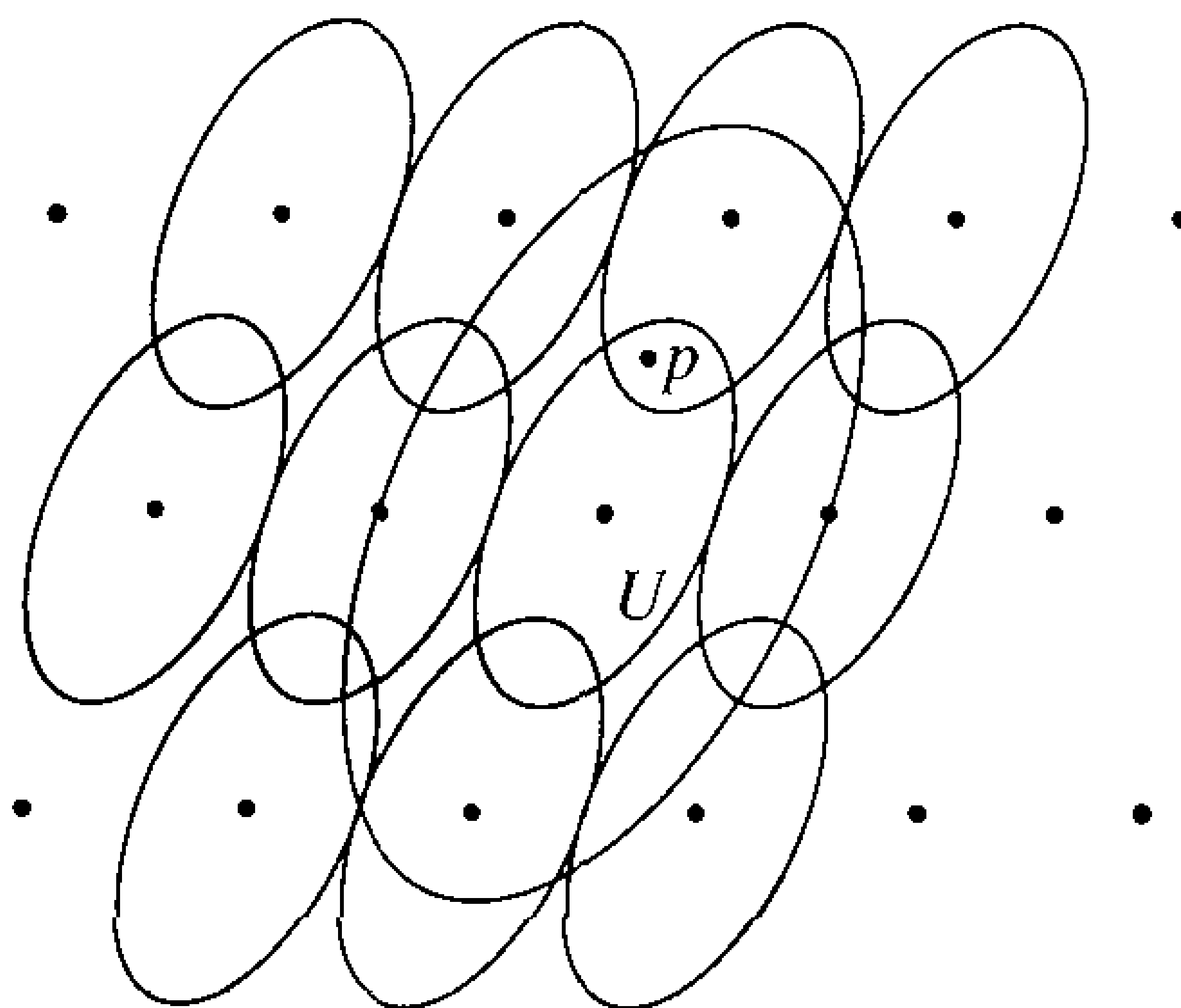
Notice that these conditions imply that $0 \in S$, unless S is empty.

(10.9) **Minkowski's Lemma.** Let L be a lattice in $\mathbb{R}^2$, and let S be a convex, centrally symmetric subset of $\mathbb{R}^2$. Let $\Delta(L)$ denote the area of the parallelogram spanned by a lattice basis for L. If

$$\text{Area}(S) > 4\Delta(L),$$

then S contains a lattice point other than 0.

Proof. Define U to be the convex set similar to S, but with half the linear dimension. In other words, we put $p \in U$ if $2p \in S$. Then U is also convex and centrally symmetric, and $\text{Area}(U) = \frac{1}{4}\text{Area}(S)$. So the above inequality can be restated as $\text{Area}(U) > \Delta(L)$.



(10.10) **Figure.**

(10.11) **Lemma.** There is an element $\alpha \in L$ such that $U \cap (U + \alpha)$ is not empty.

Proof. Let P be the parallelogram spanned by a lattice basis for L. The translates $P + \alpha$ with $\alpha \in L$ cover the plane without overlapping except along their edges. The heuristic reason that the lemma is true is this: There is one translate $U + \alpha$ for each translate $P + \alpha$, and the area of U is larger than the area of P. So the translates $U + \alpha$ must overlap. To make this precise, we note that since U is a bounded set, it meets finitely many of the translates $P + \alpha$, say it meets $P + \alpha_1, \dots, P + \alpha_k$. Denote by U_i the set $(P + \alpha_i) \cap U$. Then U is cut into the pieces $U_1, \dots, U_k$, and $\text{Area}(U) = \Sigma\,\text{Area}(U_i)$. We translate U_i back to P by subtracting α_i, setting $V_i = U_i - \alpha_i$, and we note that $V_i = P \cap (U - \alpha_i)$. So V_i is a subset of P, and $\text{Area}(V_i) = \text{Area}(U_i)$. Then $\Sigma\,\text{Area}(V_i) = \text{Area}(U) > \Delta(L) = \text{Area}(P)$. This implies that two of the sets V_i must overlap, that is, that for some $i \neq j$, $(U - \alpha_i) \cap (U - \alpha_j)$ is nonempty. Adding α_i and setting $\alpha = \alpha_i - \alpha_j$, we find that $U \cap (U + \alpha)$ is nonempty too.

Returning to the proof of Minkowski's Lemma, choose α as in Lemma (10.11), and let p be a point of $U \cap (U + \alpha)$. From $p \in U + \alpha$, it follows that $p - \alpha \in U$. By central symmetry, $q = \alpha - p \in U$ too. The midpoint between p and q is $\frac{1}{2}\alpha$, which is also in U, because U is convex. Therefore $\alpha \in S$, as required. □

(10.12) **Corollary.** Any lattice L in $\mathbb{R}^2$ contains a nonzero vector α such that

$$|\alpha|^2 \leq 4\Delta(L)/\pi.$$

Proof. We apply Minkowski's Lemma, taking for S a circle of radius r about the origin. The lemma guarantees the existence of a nonzero lattice point in S, provided that $\pi r^2 > 4\Delta(L)$, or that $r^2 > 4\Delta(L)/\pi$. So for any positive number ϵ, there is a lattice point α with $|\alpha|^2 < 4\Delta(L)/\pi + \epsilon$. Since there are only finitely many lattice points in a bounded region and since ϵ can be arbitrarily small, there is a lattice point satisfying the desired inequality. □

We now return to ideals in the ring R of integers in an imaginary quadratic field. There are two measures for the size of an ideal, which turn out to be the same. The first is the index in R. Since an ideal A is a sublattice of R, it has finite index:

$$[R : A] = \textit{number of additive cosets of } A \textit{ in } R.$$

This index can be expressed in terms of the area of the parallelogram spanned by basis vectors:

(10.13) **Lemma.** Let (a_1, a_2) and (b_1, b_2) be lattice bases for lattices $B \supset A$ in $\mathbb{R}^2$, and let $\Delta(A)$ and $\Delta(B)$ be the areas of the parallelograms spanned by these bases. Then $[B : A] = \Delta(A)/\Delta(B)$.

We leave the proof as an exercise. □

(10.14) **Corollary.**

(a) Let A be a plane lattice. The area $\Delta(A)$ is independent of the lattice basis for A.
(b) If $C \supset B \supset A$ are lattices, then $[C : A] = [C : B][B : A]$. □

It is easy to compute the area $\Delta(R)$ using the description (6.14) of the ring:

$$(10.15) \qquad \Delta(R) = \tfrac{1}{2}\sqrt{|D|} = \begin{cases} \sqrt{|d|} & \text{if } d \equiv 2, 3 \pmod 4 \\ \tfrac{1}{2}\sqrt{|d|} & \text{if } d \equiv 1 \pmod 4 \end{cases},$$

where D is the discriminant (6.18).

The other measure of the size of an ideal can be obtained from the Main Lemma (8.10): We write $A\bar{A} = (n)$ and take the integer n (chosen > 0, of course). This is analogous to the norm of an element (7.1) and is therefore called the *norm* of the ideal:

$$(10.16) \qquad N(A) = n, \quad \text{if } A\bar{A} = (n).$$

It has the multiplicative property

$$(10.17) \qquad N(AB) = N(A)N(B),$$

because $AB\overline{AB} = A\bar{A}B\bar{B} = (nm)$ if $N(B) = m$. Note also that if A is the principal ideal (α), then its norm is the norm of α:

$$(10.18) \qquad N((\alpha)) = \alpha\bar{\alpha} = N(\alpha),$$

because $(\alpha)(\bar{\alpha}) = (\alpha\bar{\alpha})$.

(10.19) **Lemma.** For any nonzero ideal A of R,

$$[R : A] = N(A).$$

(10.20) **Corollary.** *Multiplicative property of the index:* Let A and B be nonzero ideals of R. Then

$$[R : AB] = [R : A][R : B]. \square$$

Let us defer the proof of Lemma (10.19) and derive the finiteness of the class number from it.

(10.21) **Theorem.** Let $\mu = 2\sqrt{|D|}/\pi$. Every ideal class contains an ideal A such that $N(A) \leq \mu$.

Proof. Let A be an ideal. We have to find another ideal A' in the class of A whose norm is not greater than μ. We apply Corollary (10.12): There is an element $\alpha \in A$ with

$$N(\alpha) = |\alpha|^2 \leq 4\Delta(A)/\pi.$$

Then $A \supset (\alpha)$. This implies that A divides (α), that is, that $AC = (\alpha)$ for some ideal C. By the multiplicative property of norms (10.17) and by (10.18), $N(A)N(C) = N(\alpha) \leq 4\Delta(A)/\pi$. Using (10.13), (10.14), and (10.19), we write $\Delta(A) = [R:A]\Delta(R) = \frac{1}{2}N(A)\sqrt{|D|}$. Substituting for $\Delta(A)$ and cancelling $N(A)$, we find $N(C) \leq \mu$.

Now since CA is a principal ideal, the class $\langle C \rangle$ is the inverse of $\langle A \rangle$, i.e., $\langle C \rangle = \langle \overline{A} \rangle$. So we have shown that $\langle \overline{A} \rangle$ contains an ideal whose norm satisfies the required inequality. Interchanging the roles of A and $\overline{A}$ completes the proof. $\square$

The finiteness of the class number follows easily:

(10.22) **Theorem.** The ideal class group $\mathcal{C}$ is finite.

Proof. Because of (10.19) and (10.21), it is enough to show that there are finitely many ideals with index $[R:A] \leq \mu$, so it is enough to show that there are only finitely many sublattices $L \subset R$ with $[R : L] \leq \mu$. Choose an integer $n \leq \mu$, and let L be a sublattice such that $[R : L] = n$. Then R/L is an abelian group of order n, so multiplication by n is the zero map on this group. The translation of this fact to R is the statement $nR \subset L$: Sublattices of index n contain nR. Lemma (8.7) implies that there are finitely many such lattices L. Since there are also finitely many possibilities for n, we are done. $\square$

The ideal class group can be computed explicitly by checking which of the sublattices $L \subset R$ of index $\leq \mu$ are ideals. However, this is not efficient. It is better to look directly for prime ideals. Let $[\mu]$ denote the largest integer less than μ.

(10.23) **Proposition.** The ideal class group $\mathscr{C}$ is generated by the classes of the prime ideals P which divide integer primes $p \leq [\mu]$.

Proof. We know that every class contains an ideal A of norm $N(A) \leq \mu$, and since $N(A)$ is an integer, $N(A) \leq [\mu]$. Suppose that an ideal A with norm $\leq \mu$ is factored into prime ideals: $A = P_1 \cdots P_r$. Then $N(A) = N(P_1) \cdots N(P_k)$, by (10.17). Hence $N(P_i) \leq [\mu]$ for each i. So the classes of prime ideals P of norm $\leq [\mu]$ form a set of generators of $\mathscr{C}$, as claimed. □

To apply this proposition, we examine each prime integer $p \leq [\mu]$. If p remains prime in R, then the prime ideal (p) is principal, so its class is trivial. We throw out these primes. If p does not remain prime in R, then we include the class of one of its two prime ideal factors P in our set of generators. The class of the other prime factor is its inverse. It may still happen that P is a principal ideal, in which case we discard it. The remaining primes generate $\mathscr{C}$.

Table (10.24) gives a few values which illustrate different groups.

TABLE 10.24 SOME IDEAL CLASS GROUPS

d	D	$[\mu]$	*Ideal class group*
−2	−8	1	trivial
−5	−20	2	order 2
−13	−52	4	order 2
−14	−56	4	order 4, cyclic
−21	−84	5	Klein four group
−23	−23	3	order 3
−26	−104	6	order 6
−47	−47	4	order 5
−71	−71	5	order 7

(10.25) **Examples.** To apply Proposition (10.23), we factor (p) into prime ideals for all prime integers $p \leq \mu$.

(a) $d = -7$. In this case $[\mu] = 1$. Proposition (10.23) tells us that the class group $\mathscr{C}$ is generated by the empty set of prime ideals. So $\mathscr{C}$ is trivial, and R is a unique factorization domain.

(b) $d = -67$. Here $R = \mathbb{Z}[\eta]$, where $\eta = \frac{1}{2}(1 + \delta)$, and $[\mu] = 5$. The ideal class group is generated by the prime ideals dividing $2, 3, 5$. According to Proposition (9.3), a prime integer p remains prime in R if and only if the polynomial $x^2 - x + 17$ is irreducible modulo p. This is true for each of the primes $2, 3, 5$. So the primes in question are principal, and the ideal class group is trivial.

(c) $d = -14$. Here $[\mu] = 4$, so $\mathscr{C}$ is generated by prime ideals dividing (2) and (3). The polynomial $x^2 + 14$ is reducible, both modulo 2 and modulo 3, so by (9.3) neither of these integers remains prime in R. Say that $(2) = P\overline{P}$ and $(3) = Q\overline{Q}$. As in the discussion of $\mathbb{Z}[\sqrt{-5}]$, we find that $P = (2, \delta) = \overline{P}$. The ideal class $\langle P \rangle$ has order 2 in $\mathscr{C}$.

To compute the order of the class $\langle Q \rangle$, we may compute the powers of the ideal explicitly and find the first power whose lattice is similar to R. This is not efficient. It is better to compute the norms of a few small elements of R, hoping to deduce a relation among the generators. The most obvious elements to try are δ and $1 + \delta$. But $N(\delta) = 14$ and $N(1 + \delta) = 15$. These are not as good as we may hope for, because they involve the primes 7 and 5, whose factors are not among our generators. We'd rather not bring in these extra primes. The element $2 + \delta$ is better: $N(2 + \delta) = (2 + \delta)(2 - \delta) = 2 \cdot 3 \cdot 3$. This gives us the ideal relation

$$(2 + \delta)(2 - \delta) = P\overline{P}Q\overline{Q}Q\overline{Q} = P^2Q^2\overline{Q}^2.$$

Since $2 + \delta$ and $2 - \delta$ are not associates, they do not generate the same ideal. On the other hand, they generate conjugate ideals. Taking these facts into account, the only possible prime factorizations of $(2 + \delta)$ are PQ^2 and $P\overline{Q}^2$. Which case we have depends on which factor of (3) we label as Q. So we may suppose that $(2 + \delta) = PQ^2$. Then since $(2 + \delta)$ is a principal ideal, $\langle P \rangle \langle Q \rangle^2 = \underline{1}$ in $\mathscr{C}$. Hence $\langle Q \rangle^2 = \langle P \rangle^{-1} = \langle P \rangle$. This shows that $\mathscr{C}$ is the cyclic group of order 4 generated by $\langle Q \rangle$.

(d) $d = -23$, and hence $R = \mathbb{Z}[\eta]$ where $\eta = \frac{1}{2}(1 + \delta)$. Then $[\mu] = 3$, so $\mathscr{C}$ is generated by the classes of the prime ideals dividing (2) and (3). Both of these primes split in R, because the polynomial $x^2 - x + 6$ is reducible modulo 2 and modulo 3 (9.3). In fact, $(2) = P\overline{P}$, where P has the lattice base $(2,\eta)$ [see (7.8)]. This is not a principal ideal.

Say that $(3) = Q\overline{Q}$. To determine the structure of the ideal class group, we note that $N(\eta) = 2 \cdot 3$ and $N(1+\eta) = 2 \cdot 2 \cdot 2$. Therefore

$$(\eta)(\overline{\eta}) = P\overline{P}Q\overline{Q} \quad \text{and} \quad (1+\eta)(\overline{1+\eta}) = (8) = (2)^3 = P^3\overline{P}^3.$$

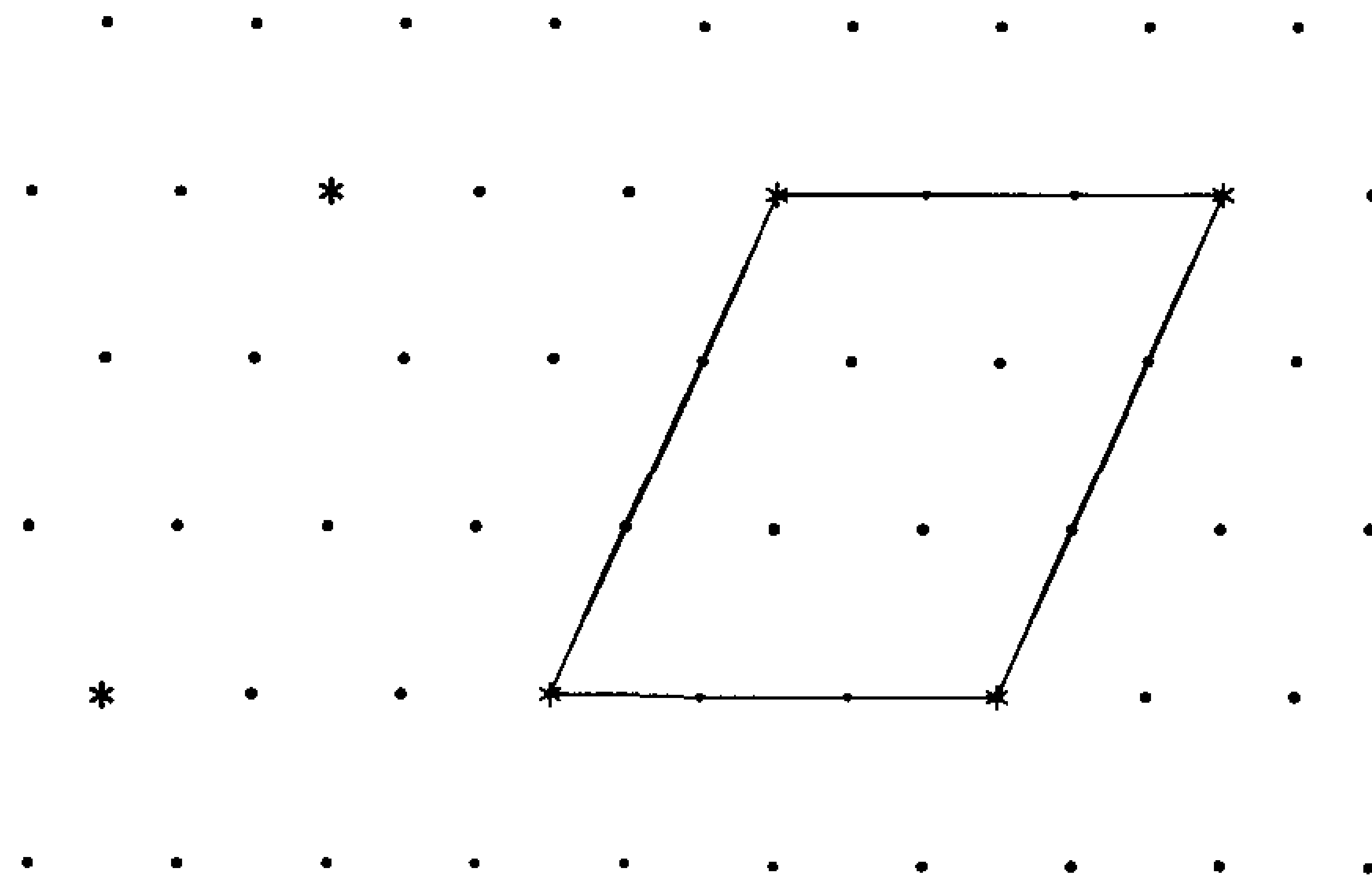
Interchanging the roles of $P, \overline{P}$ and of $Q, \overline{Q}$ as necessary, we obtain $(\eta) = PQ$ and $(1 + \eta) = P^3$ or $\overline{P}^3$. Therefore $\langle P \rangle^3 = \langle 1 \rangle$ and $\langle Q \rangle = \langle P \rangle^{-1}$ in $\mathscr{C}$. The ideal class group is a cyclic group of order 3. □

Proof of Lemma (10.19). This lemma is true for the unit ideal R. We will prove that $[R : P] = N(P)$ if P is a prime ideal, and we will show that if P is prime and if A is an arbitrary nonzero ideal, then $[R : AP] = [R : A][R : P]$. It will follow that if $[R{:}A] = N(A)$, then $[R{:}AP] = N(AP)$. Induction on the length of the prime factorization of an ideal will complete the proof.

(10.26) **Lemma.** Let n be an ordinary integer, and let A be an ideal. Then

$$[R : nA] = n^2[R : A].$$

Proof. We know that $R \supset A \supset nA$, and therefore (10.14b) $[R : nA] = [R : A][A : nA]$. Thus we must show that $[A : nA] = n^2$. Now A is a lattice, and nA is the sublattice obtained by stretching by the factor n:



(10.27) **Figure**. $3A = \{*\}$.

Clearly, $[A : nA] = n^2$, as required. □

We return to the proof of Lemma (10.19). There are two cases to consider for the ideal P. According to (9.1), there is an integer prime p so that either $P = (p)$ or $P\overline{P} = (p)$.

In the first case, $N(P) = p^2$, and $AP = pA$. We can use Lemma (10.26) twice to conclude that $[R : AP] = p^2[R : A]$ and $[R : P] = p^2[R : R] = p^2$. Thus $[R : AP] = [R : A][R : P]$ and $[R : P] = N(P)$, as required.

In the second case, $N(P) = p$. We consider the chain of ideals $A > AP > AP\overline{P}$. It follows from the Cancellation Law (8.11a) that this is a strictly decreasing chain of ideals, hence that

$$(10.28) \qquad [R : A] < [R : AP] < [R : AP\overline{P}].$$

Also, since $P\overline{P} = (p)$, we have $AP\overline{P} = pA$. Therefore we may apply Lemma (10.26) again, to conclude that $[R : AP\overline{P}] = p^2[R : A]$. Since each index (10.28) is a proper division of the next, the only possibility is that $[R : AP] = p[R : A]$. Applying this to the case $A = R$ shows that $[R : P] = p = N(P)$. So we find $[R : AP] = [R : A][R : P]$ and $[R : P] = N(P)$ again. This completes the proof. □

11. REAL QUADRATIC FIELDS

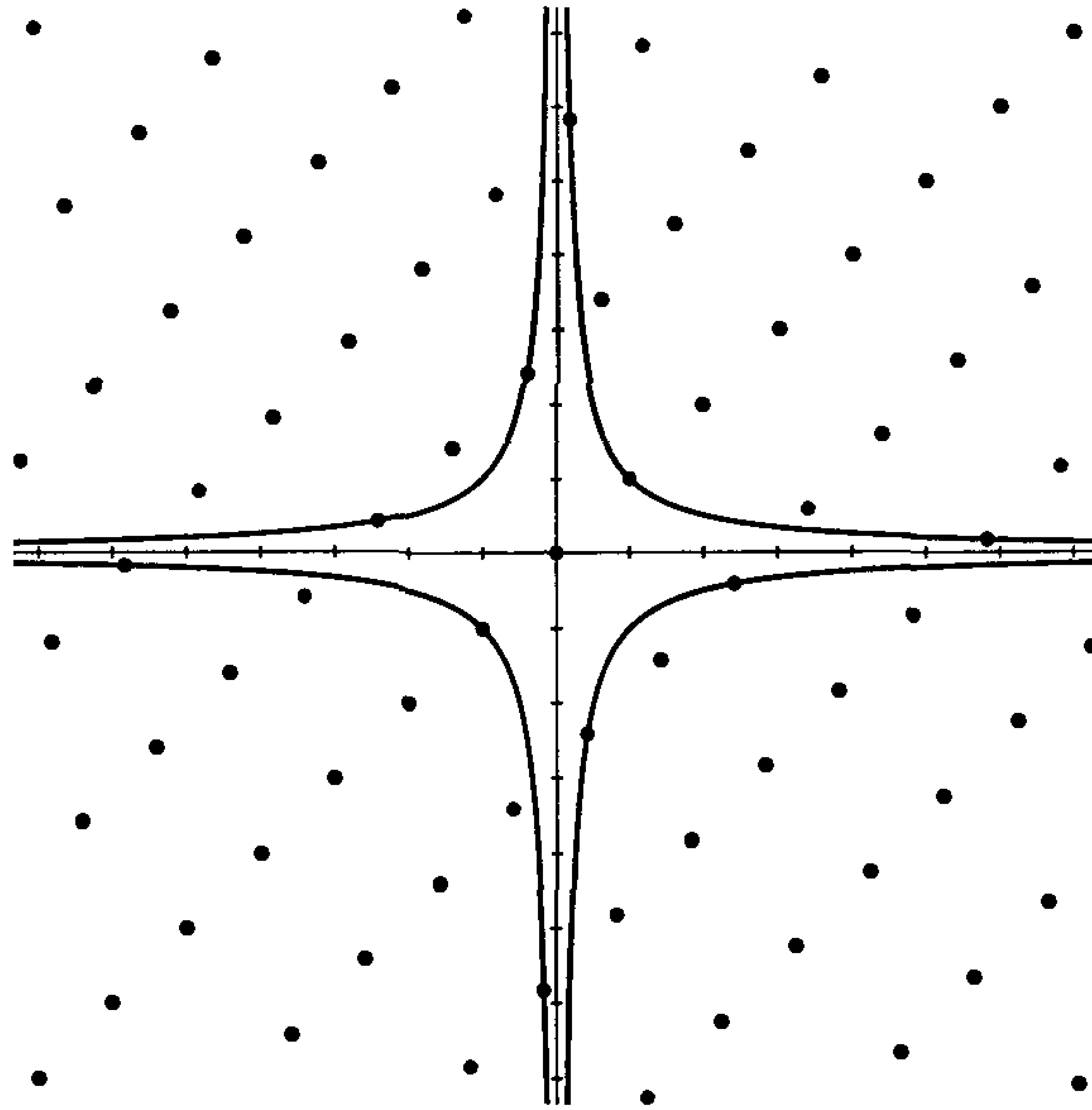
In this section we will take a brief look at real quadratic number fields $\mathbb{Q}[\delta]$, where $\delta^2 = d > 0$. We will use the field $\mathbb{Q}[\sqrt{2}]$ as an example. The ring of integers in this field is

$$(11.1) \qquad R = \mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Z}\}.$$

Since $Q[\sqrt{d}]$ is a subfield of the real numbers, the ring of integers is not embedded as a lattice in the complex plane, but we can represent R as a lattice by using the coefficients (a, b) as coordinates. A slightly more convenient representation of R as a lattice is obtained by associating to the algebraic integer $a + b\sqrt{d}$ the point (u, v), where

$$u = a + b\sqrt{d}, \quad v = a - b\sqrt{d}. \tag{11.2}$$

The resulting lattice is depicted below for the case $d = 2$:



(11.3) **Figure.** The lattice $\mathbb{Z}[\sqrt{2}]$.

Since the (u, v)-coordinates are related to the (a, b)-coordinates by the linear transformation (11.2), there is no essential difference between the two ways of depicting R, though since the transformation is not orthogonal, the shape of the lattice is different in the two representations.

Recall that the field $\mathbb{Q}[\sqrt{d}]$ is isomorphic to the abstractly constructed field

$$F \approx \mathbb{Q}[x]/(x^2 - d). \tag{11.4}$$

Let us replace $\mathbb{Q}[\sqrt{d}]$ by F and denote the residue of x in F by δ. So this element δ is an abstract square root of d rather than the positive real square root. Then the coordinates u, v represent the two ways that the abstractly given field F can be embedded into the real numbers; namely u sends $\delta \rightsquigarrow \sqrt{d}$ and v sends $\delta \rightsquigarrow -\sqrt{d}$.

For $\alpha = a + b\delta \in \mathbb{Q}[\delta]$, let us denote by α' the "conjugate" element $a - b\delta$. The *norm* of α is defined to be

$$N(\alpha) = \alpha\alpha' = a^2 - db^2, \tag{11.5}$$

in analogy with the imaginary quadratic case (7.1). If α is an algebraic integer, then $N(\alpha)$ is an integer, not necessarily positive, and

$$N(\alpha\beta) = \alpha\beta\alpha'\beta' = N(\alpha)N(\beta). \tag{11.6}$$

With this definition of norm, the proof of unique factorization of ideals into prime ideals in imaginary quadratic fields carries over.

There are two notable differences between real and imaginary quadratic fields. The first is that, for real quadratic fields, ideals in the same class are not similar geometric figures when embedded as lattices in the (u, v)-plane by (11.2). In particular, principal ideals need not be similar to the lattice R. The reason is simple: Multiplication by an element $\alpha = a + b\delta$ stretches the u-coordinate by the factor $a + b\sqrt{d}$, and it stretches the v-coordinate by the different factor $a - b\sqrt{d}$. This fact complicates the geometry slightly, and it is the reason that we developed the imaginary quadratic case first. It does not change the theory in an essential way: The class number is still finite.

The second difference is more important. It is that there are infinitely many units in the rings of integers in a real quadratic field. Since the norm $N(\alpha)$ of an algebraic integer is an ordinary integer, a unit must have norm ± 1 as before [see (7.3)], and if $N(\alpha) = \alpha\alpha' = \pm 1$, then $\pm\alpha'$ is the inverse of α, so α is a unit. For example,

$$\alpha = 1 + \sqrt{2}, \quad \alpha^2 = 3 + 2\sqrt{2} \tag{11.7}$$

are units in the ring $R = \mathbb{Z}[\sqrt{2}]$. Their norms are -1 and 1 respectively. The element α has infinite order in the group of units of R.

The condition $N(\alpha) = a^2 - 2b^2 = \pm 1$ for units translates in (u, v)-coordinates to

$$uv = \pm 1. \tag{11.8}$$

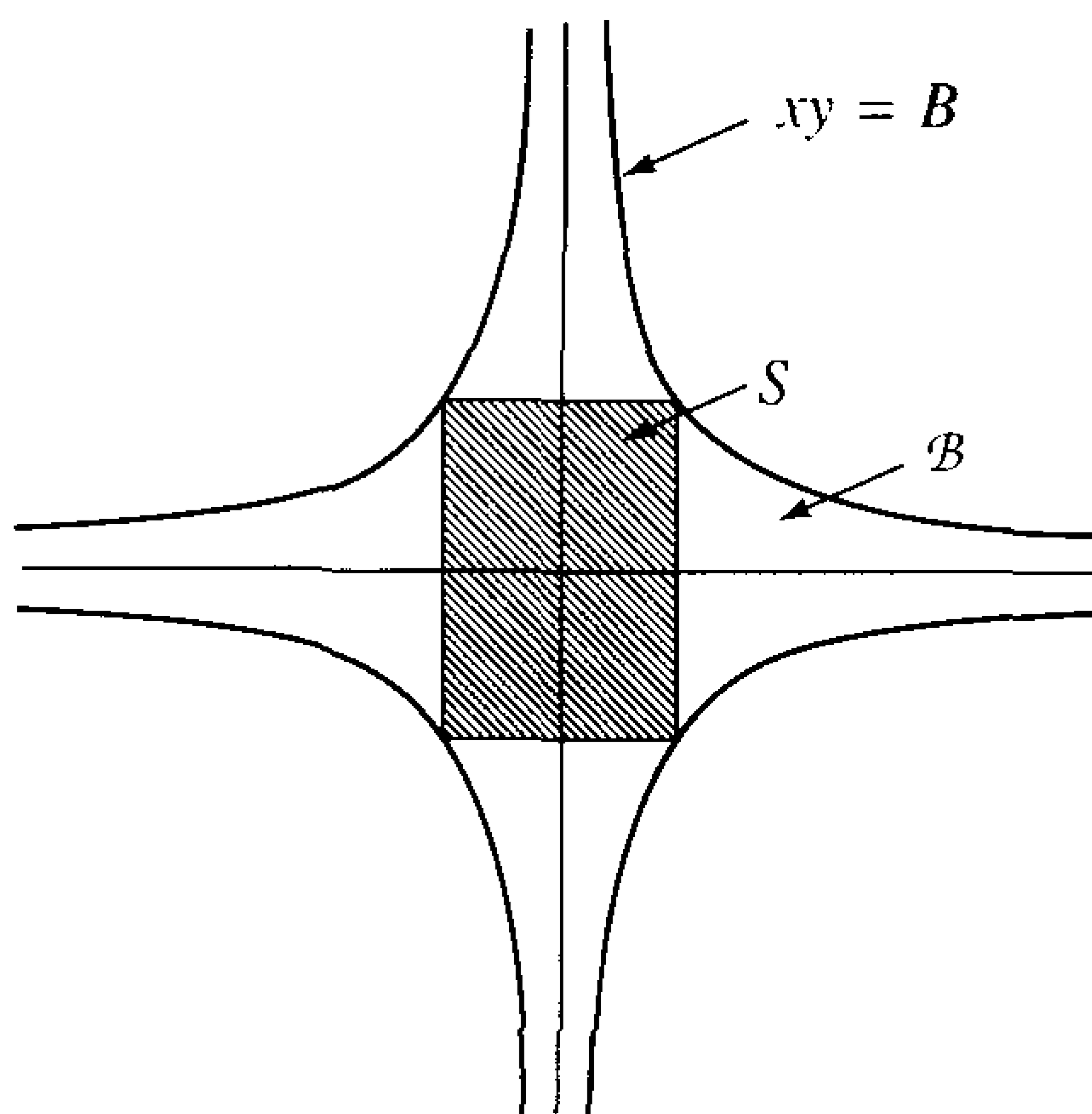
The units are the points of the lattice which lie on one of the two hyperbolas $uv = 1$ and $uv = -1$. These hyperbolas are depicted in Figure (11.3). It is a remarkable fact that real quadratic fields always have infinitely many units or, what amounts to the same thing, that the lattice of integers always contains infinitely many points on the hyperbola $uv = 1$. This fact is not obvious, either algebraically or geometrically.

(11.9) **Theorem.** Let R be the ring of integers in a real quadratic number field. The group of units in R is infinite.

(11.10) **Lemma.** Let Δ denote the area of the parallelogram spanned by a lattice basis of R, in its embedding into the (u, v)-plane. There are infinitely many elements β of R whose norm $N(\beta)$ is bounded, in fact, such that $|N(\beta)| \leq B$, where B is any real number $> \Delta$.

Proof. In the embedding into the (u, v)-plane, the elements of norm r are the lattice points on the hyperbola $xy = r$, and the elements whose norm is bounded in

absolute value by a positive number B are those lying in the region $\mathcal{B}$ bounded by the four branches of the hyperbolas $xy = B, xy = -B$.



(11.11) **Figure.**

Choose an arbitrary positive real number u_0. Then the rectangle S whose vertices are $(\pm u_0, \pm B/u_0)$ lies entirely in the region $\mathcal{B}$, and the area of this rectangle is $4B$. So if $B > \Delta$, then Minkowski's Lemma guarantees the existence of a nonzero lattice point α in S. The norm of this point is bounded by B. This is true for all u_0, and if u_0 is very large, the rectangle S is very narrow. On the other hand, there are no lattice points on the u_0-axis, because there are no nonzero elements in R of norm zero. So no particular lattice point is contained in all the rectangles S. It follows that there are infinitely many lattice points in $\mathcal{B}$. □

Since there are only finitely many integers r in the interval $-B \leq r \leq B$, Lemma (11.10) implies the following corollary:

(11.12) **Corollary.** For some integer r, there are infinitely many elements of R of norm r. □

Let r be an integer. We will call two elements $\beta_i = m_i + n_i\delta$ of R *congruent modulo* r if r divides $\beta_1 - \beta_2$ in R. If $d \equiv 2$ or 3 (modulo 4), this just means that $m_1 \equiv m_2$ and $n_1 \equiv n_2$ (modulo r).

(11.13) **Lemma.** Let β_1, β_2 be elements of R with the same norm r, and which are congruent modulo r. Then β_1/β_2 is a unit of R.

Proof. It suffices to show that β_1/β_2 is in R, because the same argument will show that $\beta_2/\beta_1 \in R$, hence that β_1/β_2 is a unit. Let $\beta_i' = m_i - n_i\delta$ be the conjugate of β_i. Then $\beta_1/\beta_2 = \beta_1\beta_2'/\beta_2\beta_2' = \beta_1\beta_2'/r$. But $\beta_2' \equiv \beta_1'$ (modulo r), so $\beta_1\beta_2' \equiv \beta_1\beta_1' = r$ (modulo r). Therefore r divides $\beta_1\beta_2'$, which shows that $\beta_1/\beta_2 \in R$, as required. □

Proof of Theorem (11.9). We choose r so that there are infinitely many elements $\beta = m + n\delta$ of norm r. We partition the set of these elements according to the congruence classes modulo r. Since there are finitely many congruence classes, some class contains infinitely many elements. The ratio of any two of these elements is a unit. □

12. SOME DIOPHANTINE EQUATIONS

Diophantine equations are polynomial equations with integer coefficients, which are to be solved in the integers. The most famous is the *Fermat Equation*

$$x^n + y^n = z^n. \tag{12.1}$$

Fermat's "Last Theorem" asserts that if $n \geq 3$ this equation has no integer solutions x, y, z, except for the trivial solutions in which one of the variables is zero. Fermat wrote this theorem in the margin of a book, asserting that the margin did not contain enough room for his proof. No proof is known today, though the theorem has been proved for all $n < 10^5$. Also, a theorem proved by Faltings in 1983, which applies to this equation as well as to many others, shows that there are only *finitely many* integer solutions for any given value of n.

This section contains a few examples of Diophantine equations which can be solved using the arithmetic of imaginary quadratic numbers. They are included only as samples. An interested reader should look in a book on number theory for a more organized discussion.

We have two methods at our disposal, namely arithmetic of quadratic number fields and congruences, and we will use both.

(12.2) **Example.** Determination of the integers n such that the equation

$$x^2 + y^2 = n$$

has an integer solution.

Here the problem is to determine the integers n which are sums of two squares or, equivalently, such that there is a point with integer coordinates on the circle $x^2 + y^2 = n$. Theorem (5.1) tells us that when p is a prime, the equation $x^2 + y^2 = p$ has an integer solution if and only if either $p = 2$ or $p \equiv 1$ (modulo 4). It is not difficult to extend this result to arbitrary integers. To do so, we interpret a sum of squares $a^2 + b^2$ as the norm $\alpha\overline{\alpha}$ of the Gauss integer $\alpha = a + bi$. Then the problem is to decide which integers n are the norms of Gauss integers. Now if a Gauss integer α is factored into Gauss primes, say $\alpha = \pi_1 \cdots \pi_k$, then its norm factors too: $N(\alpha) = N(\pi_1) \cdots N(\pi_k)$. So if n is the norm of a Gauss integer, then it is a product of norms of Gauss primes, and conversely. The norms of Gauss primes are the primes $p \equiv 1$ (modulo 4), the squares of primes $p \equiv 3$ (modulo 4), and the prime 2. Thus we have the following theorem:

(12.3) **Theorem.** The equation $x^2 + y^2 = n$ has an integer solution if and only if every prime p which is congruent 3 modulo 4 has an even exponent in the factorization of n. □

(12.4) **Example.** Determination of the integer solutions of the equation

$$y^2 + 13 = x^3.$$

We factor the left side of the equation, obtaining

$$(y + \delta)(y - \delta) = x^3,$$

where $\delta = \sqrt{-13}$. The ring of integers $R = \mathbb{Z}[\delta]$ is not a unique factorization domain, so we will analyze this equation using ideal factorization.

(12.5) **Lemma.** Let a, b be integers, and let R be any ring containing $\mathbb{Z}$ as a subring. If a and b are contained in a common proper ideal A of R, then they have a common prime factor in $\mathbb{Z}$.

Proof. We prove the contrapositive. If a, b have no common prime factor in $\mathbb{Z}$, then we can write $1 = ra + sb$, $r, s \in \mathbb{Z}$. This equation shows that if a, b are in an ideal A of R, then $1 \in A$ too. Hence A is not a proper ideal. □

(12.6) **Lemma.** Let x, y be an integer solution of the equation (12.4). The two elements $y + \delta$ and $y - \delta$ have no common prime ideal factor in R.

Proof. Let P be a prime ideal of R which contains $y + \delta$ and $y - \delta$. Then $2y \in P$ and $2\delta \in P$. Since P is a prime ideal, either $2 \in P$, or else $y \in P$ and $\delta \in P$.

In the first case, 2 and $y^2 + 13$ are not relatively prime integers by Lemma (12.5), and since 2 is prime, it divides $y^2 + 13$ in $\mathbb{Z}$. This implies that 2 divides x and that 8 divides $y^2 + 13 = x^3$. So y must be odd. Then $y^2 \equiv 1$ (modulo 4); hence $y^2 + 13 \equiv 2$ (modulo 4). This contradicts $x^3 \equiv 0$ (modulo 8).

Suppose that $y, \delta \in P$. Then $13 \in P$, and hence 13 and y are not relatively prime in $\mathbb{Z}$, that is, 13 divides y. Therefore 13 divides x, and reading the equation $y^2 + 13 = x^3$ modulo 13^2, we obtain $13 \equiv 0$ (modulo 13^2), which is a contradiction. So we have shown that $y + \delta$ and $y - \delta$ are relatively prime in R. □

We now read the equation $(y + \delta)(y - \delta) = (x)^3$ as an equality of principal ideals of R, and we factor the right side into primes, say

$$(y + \delta)(y - \delta) = (P_1 \cdots P_s)^3.$$

On the right we have a cube, and the two ideals on the left have no common prime factor. It follows that each of these ideals is a cube too, say $(y + \delta) = A^3$ and $(y - \delta) = \overline{A}^3$ for some ideal A. Looking at our table of ideal classes, we find that the ideal class group of R is cyclic of order 2. So the ideal classes of A and A^3 are equal. Since A^3 is a principal ideal, so is A, say $A = (u + v\delta)$, for some integers

u, v. We have been lucky. Since the units in R are ± 1, $(u + v\delta)^3 = \pm(y + \delta)$. Changing sign if necessary, we may assume that $(u + v\delta)^3 = y + \delta$.

We now complete the analysis by studying the equation $y + \delta = (u + v\delta)^3$. We expand the right side, obtaining

$$y + \delta = (u^3 - 39uv^2) + (3u^2v - 13v^3)\delta.$$

So $y = u^3 - 39uv^2$ and $1 = (3u^2 - 13v^2)v$. The second equation implies that $v = \pm 1$ and that $3u^2 - 13 = \pm 1$. The only possibilities are $u = \pm 2$ and $v = -1$. Then $y = \pm 70$ and $x = (u + v\delta)(u - v\delta) = 17$. These values do give solutions, so the integer solutions of the equation $y^2 + 13 = x^3$ are $x = 17$ and $y = \pm 70$. □

(12.7) **Example.** Determination of the prime integers p such that

$$x^2 + 5y^2 = p$$

has an integer solution.

Let $\delta = \sqrt{-5}$, and let $R = \mathbb{Z}[\delta]$. We know (9.3a) that the principal ideal (p) splits in R if and only if the congruence $x^2 \equiv -5$ (modulo p) has an integer solution. If $(p) = P\overline{P}$ and if P is a principal ideal, say $P = (a + b\delta)$, then $(p) = (a + b\delta)(a - b\delta) = (a^2 + 5b^2)$. Since the only units in R are ± 1, $a^2 + 5b^2 = \pm p$, and since $a^2 + 5b^2$ is positive, $a^2 + 5b^2 = p$.

Unfortunately, R is not a principal ideal domain. So it is quite likely that $(p) = P\overline{P}$ but that P is not a principal ideal. To analyze the situation further, we use the fact that there are exactly two ideal classes in R. The principal ideals form one class, and the other class is represented by any nonprincipal ideal. The ideal $A = (2, 1 + \delta)$ is one nonprincipal ideal, and we recall that for this ideal $A^2 = A\overline{A} = (2)$. Now since the ideal class group is cyclic of order 2, the product of any two ideals in the same class is principal. Suppose that $(p) = P\overline{P}$ and that P is not a principal ideal. Then AP is principal, say $AP = (a + b\delta)$. Then $(a + b\delta)(a - b\delta) = AP\overline{AP} = (2p)$. We find that $a^2 + 5b^2 = 2p$.

(12.8) **Lemma.** Let p be an odd prime. The congruence $x^2 \equiv -5$ (modulo p) has a solution if and only if one of the two equations $x^2 + 5y^2 = p$ or $x^2 + 5y^2 = 2p$ has an integer solution.

Proof. If the congruence has a solution, then $(p) = P\overline{P}$, and the two cases are decided as above, according to whether or not P is principal. Conversely, if $a^2 + 5b^2 = p$, then (p) splits in R, and we can apply (9.3a). If $a^2 + 5b^2 = 2p$, then $(a + b\delta)(a - b\delta) = (2p) = A\overline{A}(p)$. It follows from unique factorization of ideals that (p) splits too, so (9.3a) can be applied again. □

This lemma does not solve our original problem, but we have made progress. In most such situations we could not complete our analysis. But here we are lucky again, or rather this example was chosen because it admits a complete solution: The two cases can be distinguished by congruences. If $a^2 + 5b^2 = p$, then one of the

two integers a, b is odd and the other is even. We compute the congruence modulo 4, finding that $a^2 + 5b^2 \equiv 1$ (modulo 4). Hence $p \equiv 1$ (modulo 4) in this case. If $a^2 + 5b^2 = 2p$, we compute the congruences modulo 8. Since $p \equiv 1$ or 3 (modulo 4), we know that $2p \equiv 2$ or 6 (modulo 8). Any square is congruent 0, 1, or 4 (modulo 8). Hence $5b^2 \equiv 0, 5$, or 4 (modulo 8), which shows that $a^2 + 5b^2$ can not be congruent to 2 (modulo 8). Thus $p \equiv 3$ (modulo 4) in this case. We have therefore proved the following lemma:

(12.9) **Lemma.** Let p be an odd prime. Assume that the congruence $x^2 \equiv -5$ (modulo p) has a solution. Then $x^2 + 5y^2 = p$ has an integer solution if $p \equiv 1$ (modulo 4), and $x^2 + 5y^2 = 2p$ has an integer solution if $p \equiv 3$ (modulo 4).

There remains finally the problem of characterizing the odd primes p such that the congruence $x^2 \equiv -5$ has a solution modulo p. This is done by means of the amazing *Quadratic Reciprocity Law*, which asserts that $x^2 \equiv 5$ (modulo p) has a solution if and only if $x^2 \equiv p$ (modulo 5) has one! And the second congruence has a solution if and only if $p \equiv \pm 1$ (modulo 5). Combining this with the previous lemma and with the fact that -1 is a square modulo 5, we find:

(12.10) **Theorem.** Let p be an odd prime. The equation $x^2 + 5y^2 = p$ has an integer solution if and only if $p \equiv 1$ (modulo 4) and $p \equiv \pm 1$ (modulo 5). □

Nullum vero dubium nobis esse videtur,
quin multa eaque egregia in hoc genere adhuc lateant
in quibus alii vires suas exercere possint.

Karl Friedrich Gauss

EXERCISES

1. Factorization of Integers and Polynomials

1. Let a, b be positive integers whose sum is a prime p. Prove that their greatest common divisor is 1.
2. Define the greatest common divisor of a set of n integers, and prove its existence.
3. Prove that if d is the greatest common divisor of $a_1, \ldots, a_n$, then the greatest common divisor of $a_1/d, \ldots, a_n/d$ is 1.
4. (a) Prove that if n is a positive integer which is not a square of an integer, then $\sqrt{n}$ is not a rational number.
   (b) Prove the analogous statement for nth roots.
5. (a) Let a, b be integers with $a \neq 0$, and write $b = aq + r$, where $0 \leq r < |a|$. Prove that the two greatest common divisors (a, b) and (a, r) are equal.
   (b) Describe an algorithm, based on (a), for computing the greatest common divisor.

(c) Use your algorithm to compute the greatest common divisors of the following: (a) 1456, 235, (b) 123456789, 135792468.

6. Compute the greatest common divisor of the following polynomials: $x^3 - 6x^2 + x + 4$, $x^5 - 6x + 1$.

7. Prove that if two polynomials f, g with coefficients in a field F factor into linear factors in F, then their greatest common divisor is the product of their common linear factors.

8. Factor the following polynomials into irreducible factors in $\mathbb{F}_p[x]$.
 (a) $x^3 + x + 1, p = 2$ **(b)** $x^2 - 3x - 3, p = 5$ **(c)** $x^2 + 1, p = 7$

9. Euclid proved that there are infinitely many prime integers in the following way: If $p_1, \ldots, p_k$ are primes, then any prime factor p of $n = (p_1 \cdots p_n) + 1$ must be different from all of the p_i.
 (a) Adapt this argument to show that for any field F there are infinitely many monic irreducible polynomials in $F[x]$.
 (b) Explain why the argument fails for the formal power series ring $F[[x]]$.

10. *Partial fractions for integers:*
 (a) Write the fraction $r = 7/24$ in the form $r = a/8 + b/3$.
 (b) Prove that if $n = uv$, where u and v are relatively prime, then every fraction $r = m/n$ can be written in the form $r = a/u + b/v$.
 (c) Let $n = n_1 n_2 \cdots n_k$ be the factorization of an integer n into powers of distinct primes: $n_i = p_i^{e_i}$. Prove that every fraction $r = m/n$ can be written in the form $r = m_1/n_1 + \cdots + m_k/n_k$.

11. *Chinese Remainder Theorem:*
 (a) Let n, m be relatively prime integers, and let a, b be arbitrary integers. Prove that there is an integer x which solves the simultaneous congruence $x \equiv a$ (modulo m) and $x \equiv b$ (modulo n).
 (b) Determine all solutions of these two congruences.

12. Solve the following simultaneous congruences.
 (a) $x \equiv 3$ (modulo 15), $x \equiv 5$ (modulo 8), $x \equiv 2$ (modulo 7).
 (b) $x \equiv 13$ (modulo 43), $x \equiv 7$ (modulo 71).

13. *Partial fractions for polynomials:*
 (a) Prove that every rational function in $\mathbb{C}(x)$ can be written as sum of a polynomial and a linear combination of functions of the form $1/(x - a)^i$.
 (b) Find a basis for $\mathbb{C}(x)$ as vector space over $\mathbb{C}$.

*14. Let F be a subfield of $\mathbb{C}$, and let $f \in F[x]$ be an irreducible polynomial. Prove that f has no multiple root in $\mathbb{C}$.

15. Prove that the greatest common divisor of two polynomials f and g in $\mathbb{Q}[x]$ is also their greatest common divisor in $\mathbb{C}[x]$.

16. Let a and b be relatively prime integers. Prove that there are integers m, n such that $a^m + b^n \equiv 1$ (modulo ab).

2. Unique Factorization Domains, Principal Ideal Domains, and Euclidean Domains

1. Prove or disprove the following.
 (a) The polynomial ring $\mathbb{R}[x, y]$ in two variables is a Euclidean domain.
 (b) The ring $\mathbb{Z}[x]$ is a principal ideal domain.

2. Prove that the following rings are Euclidean domains.
(a) $\mathbb{Z}[\zeta]$, $\zeta = e^{2\pi i/3}$ (b) $\mathbb{Z}[\sqrt{-2}]$.
3. Give an example showing that division with remainder need not be unique in a Euclidean domain.
4. Let m, n be two integers. Prove that their greatest common divisor in $\mathbb{Z}$ is the same as their greatest common divisor in $\mathbb{Z}[i]$.
5. Prove that every prime element of an integral domain is irreducible.
6. Prove Proposition (2.8), that a domain R which has existence of factorizations is a unique factorization domain if and only if every irreducible element is prime.
7. Prove that in a principal ideal domain R, every pair a, b of elements, not both zero, has a greatest common divisor d, with these properties:
(i) $d = ar + bs$, for some $r, s \in R$;
(ii) d divides a and b;
(iii) if $e \in R$ divides a and b, it also divides d.
Moreover, d is determined up to unit factor.
8. Find the greatest common divisor of $(11 + 7i, 18 - i)$ in $\mathbb{Z}[i]$.
9. (a) Prove that $2, 3, 1 \pm \sqrt{-5}$ are irreducible elements of the ring $R = \mathbb{Z}[\sqrt{-5}]$ and that the units of this ring are ± 1.
(b) Prove that existence of factorizations is true for this ring.
10. Prove that the ring $\mathbb{R}[[t]]$ of formal real power series is a unique factorization domain.
11. (a) Prove that if R is an integral domain, then two elements a, b are associates if and only if they differ by a unit factor.
*(b) Give an example showing that (a) is false when R is not an integral domain.
12. Let R be a principal ideal domain.
(a) Prove that there is a *least common multiple* $[a, b] = m$ of two elements which are not both zero such that a and b divide m, and that if a, b divide an element $r \in R$, then m divides r. Prove that m is unique up to unit factor.
(b) Denote the greatest common divisor of a and b by (a, b). Prove that $(a, b)[a, b]$ is an associate of ab.
13. If a, b are integers and if a divides b in the ring of Gauss integers, then a divides b in $\mathbb{Z}$.
14. (a) Prove that the ring R (2.4) obtained by adjoining 2^k-th roots x_k of x to a polynomial ring is the union of the polynomial rings $F[x_k]$.
(b) Prove that there is no factorization of x_1 into irreducible factors in R.
15. By a *refinement* of a factorization $a = b_1 \cdots b_k$ we mean the expression for a obtained by factoring the terms b_i. Let R be the ring (2.4). Prove that any two factorizations of the same element $a \in R$ have refinements, all of whose factors are associates.
16. Let R be the ring $F[u, v, y, x_1, x_2, x_3, \ldots]/(x_1 y = uv, x_2^2 = x_1, x_3^2 = x_2, \ldots)$. Show that u, v are irreducible elements in R but that the process of factoring uv need not terminate.
17. Prove Proposition (2.9) and Corollary (2.10).
18. Prove Proposition (2.11).
19. Prove that the factorizations (2.22) are prime in $\mathbb{Z}[i]$.
20. The discussion of unique factorization involves only the multiplication law on the ring R, so it ought to be possible to extend the definitions. Let S be a commutative semigroup, meaning a set with a commutative and associative law of composition and with an iden-

tity. Suppose the Cancellation Law holds in S: If $ab = ac$ then $b = c$. Make the appropriate definitions so as to extend Proposition (2.8) to this situation.

***21.** Given elements $v_1, \ldots v_n$ in $\mathbb{Z}^2$, we can define a semigroup S as the set of all linear combinations of $(v_1, \ldots, v_n)$ with nonnegative integer coefficients, the law of composition being *addition*. Determine which of these semigroups has unique factorization.

3. Gauss's Lemma

1. Let a, b be elements of a field F, with $a \neq 0$. Prove that a polynomial $f(x) \in F[x]$ is irreducible if and only if $f(ax + b)$ is irreducible.
2. Let $F = \mathbb{C}(x)$, and let $f, g \in \mathbb{C}[x, y]$. Prove that if f and g have a common factor in $F[y]$, then they also have a common factor in $\mathbb{C}[x, y]$.
3. Let f be an irreducible polynomial in $\mathbb{C}[x, y]$, and let g be another polynomial. Prove that if the variety of zeros of g in $\mathbb{C}^2$ contains the variety of zeros of f, then f divides g.
4. Prove that two integer polynomials are relatively prime in $\mathbb{Q}[x]$ if and only if the ideal they generate in $\mathbb{Z}[x]$ contains an integer.
5. Prove Gauss's Lemma without reduction modulo p, in the following way: Let a_i be the coefficient of lowest degree i of f which is not divisible by p. So p divides a_ν if $\nu < i$, but p does not divide a_i. Similarly, let b_j be the coefficient of lowest degree of g which is not divisible by p. Prove that the coefficient of h of degree $i + j$ is not divisible by p.
6. State and prove Gauss's Lemma for Euclidean domains.
7. Prove that an integer polynomial is primitive if and only if it is not contained in any of the kernels of the maps (3.2).
8. Prove that $\det\begin{bmatrix} x & y \\ z & w \end{bmatrix}$ is irreducible in the polynomial ring $\mathbb{C}[x, y, z, w]$.
9. Prove that the kernel of the homomorphism $\mathbb{Z}[x] \longrightarrow \mathbb{R}$ sending $x \rightsquigarrow 1 + \sqrt{2}$ is a principal ideal, and find a generator for this ideal.
10. (a) Consider the map ψ: $\mathbb{C}[x, y] \longrightarrow \mathbb{C}[t]$ defined by $f(x, y) \rightsquigarrow f(t^2, t^3)$. Prove that its kernel is a principal ideal, and that its image is the set of polynomials $p(t)$ such that $p'(0) = 0$.

    (b) Consider the map φ: $\mathbb{C}[x, y] \longrightarrow \mathbb{C}[t]$ defined by $f(x, y) \rightsquigarrow (t^2 - t, t^3 - t^2)$. Prove that $\ker \varphi$ is a principal ideal, and that its image is the set of polynomials $p(t)$ such that $p(0) = p(1)$. Give an intuitive explanation in terms of the geometry of the variety $\{f = 0\}$ in $\mathbb{C}^2$.

4. Explicit Factorization of Polynomials

1. Prove that the following polynomials are irreducible in $\mathbb{Q}[x]$.
   (a) $x^2 + 27x + 213$ **(b)** $x^3 + 6x + 12$ **(c)** $8x^3 - 6x + 1$ **(d)** $x^3 + 6x^2 + 7$
   (e) $x^5 - 3x^4 + 3$
2. Factor $x^5 + 5x + 5$ into irreducible factors in $\mathbb{Q}[x]$ and in $\mathbb{F}_2[x]$.
3. Factor $x^3 + x + 1$ in $\mathbb{F}_p[x]$, when $p = 2, 3, 5$.

4. Factor $x^4 + x^2 + 1$ into irreducible factors in $\mathbb{Q}[x]$.
5. Suppose that a polynomial of the form $x^4 + bx^2 + c$ is a product of two quadratic factors in $\mathbb{Q}[x]$. What can you say about the coefficients of these factors?
6. Prove that the following polynomials are irreducible.
   (a) $x^2 + x + 1$ in the field $\mathbb{F}_2$ (b) $x^2 + 1$ in $\mathbb{F}_7$ (c) $x^3 - 9$ in $\mathbb{F}_{31}$
7. Factor the following polynomials into irreducible factors in $\mathbb{Q}[x]$.
   (a) $x^3 - 3x - 2$ (b) $x^3 - 3x + 2$ (c) $x^9 - 6x^6 + 9x^3 - 3$
8. Let p be a prime integer. Prove that the polynomial $x^n - p$ is irreducible in $\mathbb{Q}[x]$.
9. Using reduction modulo 2 as an aid, factor the following polynomials in $\mathbb{Q}[x]$.
   (a) $x^2 + 2345x + 125$ (b) $x^3 + 5x^2 + 10x + 5$ (c) $x^3 + 2x^2 + 3x + 1$
   (d) $x^4 + 2x^3 + 2x^2 + 2x + 2$ (e) $x^4 + 2x^3 + 3x^2 + 2x + 1$
   (f) $x^4 + 2x^3 + x^2 + 2x + 1$ (g) $x^5 + x^4 - 4x^3 + 2x^2 + 4x + 1$
10. Let p be a prime integer, and let $f \in \mathbb{Z}[x]$ be a polynomial of degree $2n + 1$, say $f(x) = a_{2n+1}x^{2n+1} + \cdots + a_1x + a_0$. Suppose that $a_{2n+1} \not\equiv 0$ (modulo p), $a_0, a_1, \ldots, a_n \equiv 0$ (modulo p^2), $a_{n+1}, \ldots, a_{2n} \equiv 0$ (modulo p), $a_0 \not\equiv 0$ (modulo p^3). Prove that f is irreducible in $\mathbb{Q}[x]$.
11. Let p be a prime, and let $A \neq I$ be an $n \times n$ integer matrix such that $A^p = I$ but $A \neq I$. Prove that $n \geq p - 1$.
12. Determine the monic irreducible polynomials of degree 3 over $\mathbb{F}_3$.
13. Determine the monic irreducible polynomials of degree 2 over $\mathbb{F}_5$.
14. *Lagrange interpolation formula:*
   (a) Let $x_0, \ldots, x_d$ be distinct complex numbers. Determine a polynomial $p(x)$ of degree n which is zero at $x_1, \ldots, x_n$ and such that $p(x_0) = 1$.
   (b) Let $x_0, \ldots, x_d$; $y_0, \ldots, y_d$ be complex numbers, and suppose that the x_i are all different. There is a unique polynomial $g(x) \in \mathbb{C}[x]$ of degree $\leq d$, such that $g(x_i) = y_i$ for each $i = 0, \ldots, d$. Prove this by determining the polynomial g explicitly in terms of x_i, y_i.

*15. Use the Lagrange interpolation formula to give a method of finding all integer polynomial factors of an integer polynomial in a finite number of steps.

16. Let $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ be a monic polynomial with integer coefficients, and let $r \in \mathbb{Q}$ be a rational root of $f(x)$. Prove that r is an integer.
17. Prove that the polynomial $x^2 + y^2 - 1$ is irreducible by the method of undetermined coefficients, that is, by studying the equation $(ax + by + c)(a'x + b'y + c') = x^2 + y^2 - 1$, where a, b, c, a', b', c' are unknown.

5. *Primes in the Ring of Gauss Integers*

1. Prove that every Gauss prime divides exactly one integer prime.
2. Factor 30 into primes in $\mathbb{Z}[i]$.
3. Factor the following into Gauss primes.
   (a) $1 - 3i$ (b) 10 (c) $6 + 9i$
4. Make a neat drawing showing the primes in the ring of Gauss integers in a reasonable size range.
5. Let π be a Gauss prime. Prove that π and $\overline{\pi}$ are associate if and only if either π is associate to an integer prime or $\pi\overline{\pi} = 2$.

6. Let R be the ring $\mathbb{Z}[\sqrt{3}]$. Prove that a prime integer p is a prime element of R if and only if the polynomial $x^2 - 3$ is irreducible in $\mathbb{F}_p[x]$.
7. Describe the residue ring $\mathbb{Z}[i]/(p)$ in each case.
 (a) $p = 2$ **(b)** $p \equiv 1$ (modulo 4) **(c)** $p \equiv 3$ (modulo 4)

*8. Let $R = \mathbb{Z}[\zeta]$, where $\zeta = \frac{1}{2}(-1 + \sqrt{-3})$ is a complex cube root of 1. Let p be an integer prime $\neq 3$. Adapt the proof of Theorem (5.1) to prove the following.
 (a) The polynomial $x^2 + x + 1$ has a root in $\mathbb{F}_p$ if and only if $p \equiv 1$ (modulo 3).
 (b) (p) is a prime ideal of R if and only if $p \equiv -1$ (modulo 3).
 (c) p factors in R if and only if it can be written in the form $p = a^2 + ab + b^2$, for some integers a, b.
 (d) Make a drawing showing the primes of absolute value ≤ 10 in R.

6. Algebraic Integers

1. Is $\frac{1}{2}(1 + \sqrt{3})$ an algebraic integer?
2. Let α be an algebraic integer whose monic irreducible polynomial over $\mathbb{Z}$ is $x^n + a_{n-1}n^{n-1} + \cdots + a_1x + a_0$, and let $R = \mathbb{Z}[\alpha]$. Prove that α is a unit in R if and only if $a_0 = \pm 1$.
3. Let d, d' be distinct square-free integers. Prove that $\mathbb{Q}(\sqrt{d})$ and $\mathbb{Q}(\sqrt{d'})$ are different subfields of $\mathbb{C}$.
4. Prove that existence of factorizations is true in the ring of integers in an imaginary quadratic number field.
5. Let α be the real cube root of 10, and let $\beta = a + b\alpha + c\alpha^2$, with $a, b, c, \in \mathbb{Q}$. Then β is the root of a monic cubic polynomial $f(x) \in \mathbb{Q}[x]$. The irreducible polynomial for α over $\mathbb{Q}$ is $x^3 - 10$, and its three roots are α, $\alpha' = \zeta\alpha$, and $\alpha'' = \zeta^2\alpha$, where $\zeta = e^{2\pi i/3}$. The three roots of f are β, $\beta' = a + b\zeta\alpha + c\zeta^2\alpha^2$, and $\beta'' = a + b\zeta^2\alpha + c\zeta\alpha^2$, so $f(x) = (x - \beta)(x - \beta')(x - \beta'')$.
 (a) Determine f by expanding this product. The terms involving α and α^2 have to cancel out, so they need not be computed.
 (b) Determine which elements β are algebraic integers.
6. Prove Proposition (6.17).
7. Prove that the ring of integers in an imaginary quadratic field is a maximal subring of $\mathbb{C}$ with the property of being a lattice in the complex plane.
8. **(a)** Let $S = \mathbb{Z}[\alpha]$, where α is a complex root of a monic polynomial of degree 2. Prove that S is a lattice in the complex plane.
 (b) Prove the converse: A subring S of $\mathbb{C}$ which is a lattice has the form given in (a).
9. Let R be the ring of integers in the field $\mathbb{Q}[\sqrt{d}]$.
 (a) Determine the elements $\alpha \in R$ such that $R = \mathbb{Z}[\alpha]$.
 (b) Prove that if $R = \mathbb{Z}[\alpha]$ and if α is a root of the polynomial $x^2 + bx + c$ over $\mathbb{Q}$, then the discriminant $b^2 - 4c$ is D (6.18).

7. Factorization in Imaginary Quadratic Fields

1. Prove Proposition (7.3) by arithmetic.
2. Prove that the elements $2, 3, 1 + \sqrt{-5}, 1 - \sqrt{-5}$ are irreducible elements of the ring $\mathbb{Z}[\sqrt{-5}]$.

3. Let $d = -5$. Determine whether or not the lattice of integer linear combinations of the given vectors is an ideal.
   (a) $(5, 1 + \delta)$ (b) $(7, 1 + \delta)$ (c) $(4 - 2\delta, 2 + 2\delta, 6 + 4\delta)$
4. Let A be an ideal of the ring of integers R in an imaginary quadratic field. Prove that there is a lattice basis for A one of whose elements is a positive integer.
5. Let $R = \mathbb{Z}[\sqrt{-5}]$. Prove that the lattice spanned by $(3, 1 + \sqrt{-5})$ is an ideal in R, determine its nonzero element of minimal absolute value, and verify that this ideal has the form (7.9), Case 2.
6. With the notation of (7.9), show that if α is an element of R such that $\frac{1}{2}(\alpha + \alpha\delta)$ is also in R, then $(\alpha, \frac{1}{2}(\alpha + \alpha\delta))$ is a lattice basis of an ideal.
7. For each ring R listed below, use the method of Proposition (7.9) to describe the ideals in R. Make a drawing showing the possible shapes of the lattices in each case.
   (a) $R = \mathbb{Z}[\sqrt{-3}]$ (b) $R = \mathbb{Z}[\frac{1}{2}(1 + \sqrt{-3})]$ (c) $R = \mathbb{Z}[\sqrt{-6}]$ (d) $R = \mathbb{Z}[\sqrt{-7}]$
   (e) $R = \mathbb{Z}[\frac{1}{2}(1 + \sqrt{-7})]$ (f) $R = \mathbb{Z}[\sqrt{-10}]$
8. Prove that R is not a unique factorization domain when $d \equiv 2$ (modulo 4) and $d < -2$.
9. Let $d \leq -3$. Prove that 2 is not a prime element in the ring $\mathbb{Z}[\sqrt{d}]$, but that 2 is irreducible in this ring.

8. Ideal Factorization

1. Let $R = \mathbb{Z}[\sqrt{-6}]$. Factor the ideal (6) into prime ideals explicitly.
2. Let $\delta = \sqrt{-3}$ and $R = \mathbb{Z}[\delta]$. (This is not the ring of integers in the imaginary quadratic number field $\mathbb{Q}[\delta]$.) Let A be the ideal $(2, 1 + \delta)$. Show that $A\bar{A}$ is not a principal ideal, hence that the Main Lemma is not true for this ring.
3. Let $R = \mathbb{Z}[\sqrt{-5}]$. Determine whether or not 11 is an irreducible element of R and whether or not (11) is a prime ideal in R.
4. Let $R = \mathbb{Z}[\sqrt{-6}]$. Find a lattice basis for the product ideal AB, where $A = (2, \delta)$ and $B = (3, \delta)$.
5. Prove that $A \supset A'$ implies that $AB \supset A'B$.
6. Factor the principal ideal (14) into prime ideals explicitly in $R = \mathbb{Z}[\delta]$, where $\delta = \sqrt{-5}$.
7. Let P be a prime ideal of an integral domain R, and assume that existence of factorizations is true in R. Prove that if $a \in P$ then some irreducible factor of a is in P.

9. The Relation Between Prime Ideals of R and Prime Integers

1. Find lattice bases for the prime divisors of 2 and 3 in the ring of integers in (a) $\mathbb{Q}[\sqrt{-14}]$ and (b) $\mathbb{Q}[\sqrt{-23}]$.
2. Let $d = -14$. For each of the following primes p, determine whether or not p splits or ramifies in R, and if so, determine a lattice basis for a prime ideal factor of (p): 2, 3, 5, 7, 11, 13.
3. (a) Suppose that a prime integer p remains prime in R. Prove that $R/(p)$ is then a field with p^2 elements.
   (b) Prove that if p splits in R, then $R/(p)$ is isomorphic to the product ring $\mathbb{F}_p \times \mathbb{F}_p$.

4. Let p be a prime which splits in R, say $(p) = P\overline{P}$, and let $\alpha \in P$ be any element which is not divisible by p. Prove that P is generated as an ideal by (p, α).
5. Prove Proposition (9.3b).
6. If $d \equiv 2$ or 3 (modulo 4), then according to Proposition (9.3a) a prime integer p remains prime in the ring of integers of $\mathbb{Q}[\sqrt{d}]$ if the polynomial $x^2 - d$ is irreducible modulo p.
   (a) Prove the same thing when $d \equiv 1$ (modulo 4) and $p \neq 2$.
   (b) What happens to $p = 2$ in this case?
7. Assume that $d \equiv 2$ or 3 (modulo 4). Prove that a prime integer p ramifies in R if and only if $p = 2$ or p divides d.
8. State and prove an analogue of problem 7 when d is congruent 1 modulo 4.
9. Let p be an integer prime which ramifies in R, and say that $(p) = P^2$. Find an explicit lattice basis for P. In which cases is P a principal ideal?
10. A prime integer might be of the form $a^2 + b^2d$, with $a, b \in \mathbb{Z}$. Discuss carefully how this is related to the prime factorization of (p) in R.
*11. Prove Proposition (9.1).

10. Ideal Classes in Imaginary Quadratic Fields

1. Prove that the ideals A and A' are similar if and only if there is a nonzero ideal C such that AC and $A'C$ are principal ideals.
2. The estimate of Corollary (10.12) can be improved to $|\alpha|^2 \leq 2\Delta(L)/\sqrt{3}$, by studying lattice points in a circle rather than in an arbitrary centrally symmetric convex set. Work this out.
3. Let $R = \mathbb{Z}[\delta]$, where $\delta^2 = -6$.
   (a) Prove that the lattices $P = (2, \delta)$ and $Q = (3, \delta)$ are prime ideals of R.
   (b) Factor the principal ideal (6) into prime ideals explicitly in R.
   (c) Prove that the ideal classes of P and Q are equal.
   (d) The Minkowski bound for R is $[\mu] = 3$. Using this fact, determine the ideal class group of R.
4. In each case, determine the ideal class group and draw the possible shapes of the lattices.
   (a) $d = -10$ (b) $d = -13$ (c) $d = -14$ (d) $d = -15$ (e) $d = -17$
   (f) $d = -21$
5. Prove that the values of d listed in Theorem (7.7) have unique factorization.
6. Prove Lemma (10.13).
7. Derive Corollary (10.14) from Lemma (10.13).
8. Verify Table (10.24).

11. Real Quadratic Fields

1. Let $R = \mathbb{Z}[\delta]$, $\delta = \sqrt{2}$. Define a size function on R using the lattice embedding (11.2): $\sigma(a + b\delta) = a^2 - 2b^2$. Prove that this size function makes R into a Euclidean domain.
2. Let R be the ring of integers in a real quadratic number field, with $d \equiv 2$ or 3 (modulo 4). According to (6.14), R has the form $\mathbb{Z}[x]/(x^2 - d)$. We can also consider the ring $R' = \mathbb{R}[x]/(x^2 - d)$, which contains R as a subring.
   (a) Show that the elements of R' are in bijective correspondence with points of $\mathbb{R}^2$ in such a way that the elements of R correspond to lattice points.

(b) Determine the group of units of R'. Show that the subset U' of R' consisting of the points on the two hyperbolas $xy = \pm 1$ forms a subgroup of the group of units.
(c) Show that the group of units U of R is a discrete subgroup of U', and show that the subgroup U_0 of units which are in the first quadrant is an infinite cyclic group.
(d) What are the possible structures of the group of units U?

3. Let U_0 denote the group of units of R which are in the first quadrant in the embedding (11.2). Find a generator for U_0 when (a) $d = 3$, (b) $d = 5$.
4. Prove that if d is a square >1 then the equation $x^2 - y^2 d = 1$ has no solution except $x = \pm 1, y = 0$.
5. Draw a figure showing the hyperbolas and the units in a reasonable size range for $d = 3$.

12. Some Diophantine Equations

1. Determine the primes such that $x^2 + 5y^2 = 2p$ has a solution.
2. Express the assertion of Theorem (12.10) in terms of congruence modulo 20.
3. Prove that if $x^2 \equiv -5$ (modulo p) has a solution, then there is an integer point on one of the two ellipses $x^2 + 5y^2 = p$ or $2x^2 + 2xy + 3y^2 = p$.
4. Determine the conditions on the integers a, b, c such that the linear Diophantine equation $ax + by = c$ has an integer solution, and if it does have one, find all the solutions.
5. Determine the primes p such that the equation $x^2 + 2y^2 = p$ has an integer solution.
6. Determine the primes p such that the equation $x^2 + xy + y^2 = p$ has an integer solution.
7. Prove that if the congruence $x^2 \equiv -10$ (modulo p) has a solution, then the equation $x^2 + 10y^2 = p^2$ has an integer solution. Generalize.
8. Find all integer solutions of the equation $x^2 + 2 = y^3$.
9. Solve the following Diophantine equations.
(a) $y^2 + 10 = x^3$ (b) $y^2 + 1 = x^3$ (c) $y^2 + 2 = x^3$

Miscellaneous Problems

1. Prove that there are infinitely many primes congruent 1 modulo 4.
2. Prove that there are infinitely many primes congruent to -1 (modulo 6) by studying the factorization of the integer $p_1 p_2 \cdots p_r - 1$, where $p_1, \ldots, p_r$ are the first r primes.
3. Prove that there are infinitely many primes congruent to -1 (modulo 4).
4. (a) Determine the prime ideals of the polynomial ring $\mathbb{C}[x, y]$ in two variables.
(b) Show that unique factorization of ideals does not hold in the ring $\mathbb{C}[x, y]$.
5. Relate proper factorizations of elements in an integral domain to proper factorizations of principal ideals. Using this relation, state and prove unique factorization of ideals in a principal ideal domain.
6. Let R be a domain, and let I be an ideal which is a product of distinct maximal ideals in two ways, say $I = P_1 \cdots P_r = Q_1 \cdots Q_s$. Prove that the two factorizations are the same. except for the ordering of the terms.
7. Let R be a ring containing $\mathbb{Z}$ as a subring. Prove that if integers m, n are contained in a proper ideal of R, then they have a common integer factor > 1.

***8.** (a) Let θ be an element of the group $\mathbb{R}^+/\mathbb{Z}^+$. Use the Pigeonhole Principle [Appendix (1.6)] to prove that for every integer n there is an integer $b \leq n$ such that $|b\theta| \leq 1/bn$.

(b) Show that for every real number r and every $\epsilon > 0$, there is a fraction m/n such that $|r - m/n| \leq \epsilon/n$.

(c) Extend this result to the complex numbers by showing that for every complex number α and every real number $\epsilon > 0$, there is an element of $\mathbb{Z}(i)$, say $\beta = (a + bi)/n$ with $a, b, n \in \mathbb{Z}$, such that $|\alpha - \beta| \leq \epsilon/n$.

(d) Let ϵ be a positive real number, and for each element $\beta = (a + bi)/n$ of $\mathbb{Q}(i)$, $a, b, n \in \mathbb{Z}$, consider the disc of radius ϵ/n about β. Prove that the interiors of these discs cover the complex plane.

(e) Extend the method of Proposition (7.9) to prove the finiteness of the class number for any imaginary quadratic field.

***9.** (a) Let R be the ring of functions which are polynomials in $\cos t$ and $\sin t$, with real coefficients. Prove that $R \approx \mathbb{R}[x, y]/(x^2 + y^2 - 1)$.

(b) Prove that R is not a unique factorization domain.

*(c) Prove that $\mathbb{C}[x, y]/(x^2 + y^2 - 1)$ is a principal ideal domain and hence a unique factorization domain.

***10.** In the definition of a Euclidean domain, the size function σ is assumed to have as range the set of nonnegative integers. We could generalize this by allowing the range to be some other ordered set. Consider the product ring $R = \mathbb{C}[x] \times \mathbb{C}[y]$. Show that we can define a size function $R - \{0\} \longrightarrow S$, where S is the ordered set $\{0, 1, 2, 3, \ldots; \omega, \omega + 1, \omega + 2, \omega + 3, \ldots\}$, so that the division algorithm holds.

***11.** Let φ: $\mathbb{C}[x, y] \longrightarrow \mathbb{C}[t]$ be a homomorphism, defined say by $x \rightsquigarrow x(t), y \rightsquigarrow y(t)$. Prove that if $x(t)$ and $y(t)$ are not both constant, then $\ker \varphi$ is a nonzero principal ideal.

Chapter 12

Modules

Be wise! Generalize!

Piccayune Sentinel

1. THE DEFINITION OF A MODULE

Let R be a commutative ring. An *R-module* V is an abelian group with law of composition written $+$, together with a scalar multiplication $R \times V \longrightarrow V$, written $r, v \rightsquigarrow rv$, which satisfies these axioms:

(1.1) (i) $$1v = v,$$

(ii) $$(rs)v = r(sv),$$

(iii) $$(r + s)v = rv + sv,$$

(iv) $$r(v + v') = rv + rv',$$

for all $r, s \in R$ and $v, v' \in V$. Notice that these are precisely the axioms for a vector space. An F-module is just an F-vector space, when F is a field. So modules are the natural generalizations of vector spaces to rings. But the fact that elements of a ring needn't be invertible makes modules more complicated.

The most obvious examples are the modules R^n of *R-vectors*, that is, row or column vectors with entries in the ring. The laws of composition for R-vectors are the same as for vectors with entries in a field:

$$\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1+b_1 \\ \vdots \\ a_n+b_n \end{bmatrix} \quad \text{and} \quad r\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} ra_1 \\ \vdots \\ ra_n \end{bmatrix}$$

The modules thus defined are called *free modules*. But when R is not a field, it is no longer true that these are the only modules. There will be modules which are not isomorphic to any free module, though they are spanned by a finite set.

Let us examine the concept of module in the case that R is the ring of integers $\mathbb{Z}$. Any abelian group V, its law of composition written additively, can be made into a module over $\mathbb{Z}$ in exactly one way, by the rules

$$nv = v + \cdots + v = \text{“}n \text{ times } v\text{”}$$

and $(-n)v = -(nv)$, for any positive integer n. These rules are forced on us by axioms (1.1), starting with $1v = v$, and they do make V into a $\mathbb{Z}$-module; in other words, the axioms (1.1) hold. This is intuitively very plausible. To make a formal proof, we would go back to Peano's axioms. Conversely, any $\mathbb{Z}$-module has the structure of an abelian group, given by forgetting about its scalar multiplication. Thus

(1.2) *abelian group and $\mathbb{Z}$-module are equivalent concepts.*

We must use additive notation in the abelian group in order to make this correspondence seem natural.

The ring of integers provides us with examples to show that modules over a ring need not be free. No finite abelian group except the zero group is isomorphic to a free module $\mathbb{Z}^n$, because $\mathbb{Z}^n$ is infinite if $n > 0$ and $\mathbb{Z}^0 = 0$.

The remainder of this section extends some of our basic terminology to modules. A *submodule* of an R-module V is a nonempty subset which is closed under addition and scalar multiplication. We have seen submodules in one case before, namely ideals.

(1.3) **Proposition.** The submodules of the R-module R^1 are the ideals of R.

Proof. By definition, an ideal is a subset of R which is closed under addition and under multiplication by elements of R. □

The definition of *homomorphism* of R-modules copies that of linear transformation of vector spaces. A homomorphism $\varphi: V \longrightarrow W$ of R-modules is a map which is compatible with the laws of composition

(1.4) $$\varphi(v + v') = \varphi(v) + \varphi(v') \quad \text{and} \quad \varphi(rv) = r\varphi(v),$$

for all $v, v' \in V$ and $r \in R$. A bijective homomorphism is called an *isomorphism*. The *kernel* of a homomorphism $\varphi: V \longrightarrow W$ is a submodule of V, and the *image* of φ is a submodule of W.

The proof given for vector spaces [Chapter 4 (2.1)] shows that every homomorphism $\varphi: R^m \longrightarrow R^n$ of free modules is left multiplication by a matrix whose entries are in R.

We also need to extend the concept of quotient group to modules. Let R be a ring, and let W be a submodule of an R-module V. The quotient V/W is the additive group of cosets [Chapter 2 (9.5)] $\overline{v} = v + W$. It is made into an R-module by the rule

$$r\overline{v} = \overline{rv}. \tag{1.5}$$

We have made such constructions several times before. The facts we will need are collected together below.

(1.6) **Proposition.**

(a) The rule (1.5) is well-defined, and it makes $\overline{V} = V/W$ into an R-module.
(b) The canonical map $\pi: V \longrightarrow \overline{V}$ sending $v \rightsquigarrow \overline{v}$ is a surjective homomorphism of R-modules, and its kernel is W.
(c) *Mapping property:* Let $f: V \longrightarrow V'$ be a homomorphism of R-modules whose kernel contains W. There is a unique homomorphism: $\overline{f}: \overline{V} \longrightarrow V'$ such that $f = \overline{f}\pi$.
(d) *First Isomorphism Theorem:* If $\ker f = W$, then $\overline{f}$ is an isomorphism from $\overline{V}$ to the image of f.
(e) *Correspondence Theorem:* There is a bijective correspondence between submodules $\overline{S}$ of $\overline{V}$ and submodules S of V which contain W, defined by $S = \pi^{-1}(\overline{S})$ and $\overline{S} = \pi(S)$. If S and $\overline{S}$ are corresponding modules, then V/S is isomorphic to $\overline{V}/\overline{S}$.

We already know the analogous facts for groups and normal subgroups. All that remains to be checked in each part is that scalar multiplication is well-defined, satisfies the axioms for a module, and is compatible with the maps. These verifications follow the pattern set previously. □

2. MATRICES, FREE MODULES, AND BASES

Matrices with entries in a ring can be manipulated in the same way as matrices with entries in a field. That is, the operations of matrix addition and multiplication are defined as in Chapter 1, and they satisfy similar rules. A matrix with entries in a ring R is often called an *R-matrix*.

Let us ask which R-matrices are invertible. The *determinant* of an $n \times n$ R-matrix $A = (a_{ij})$ can be computed by any of the old rules. It is convenient to use the complete expansion [Chapter 1 (4.12)], because it exhibits the determinant as a polynomial in the n^2 matrix entries. So we write

$$\det A = \sum_p \pm\, a_{1p(1)} \cdots a_{np(n)}, \tag{2.1}$$

the sum being over all permutations of the set $\{\mathbf{1}, \ldots, \mathbf{n}\}$, and the symbol $\pm$ standing for the sign of the permutation. Evaluating this formula on an R-matrix, we obtain an element of R. The usual rules for determinant apply, in particular

$$\det AB = (\det A)(\det B).$$

We have proved this rule when the matrix entries are in a field [Chapter 1 (3.16)], and we will discuss the reason that such formulas carry over to rings in the next section. Let us assume for now that they do carry over.

If A has a multiplicative inverse A^{-1} with entries in R, then

$$(\det A)(\det A^{-1}) = \det I = 1.$$

This shows that the determinant of an invertible R-matrix is a *unit* of the ring. Conversely, let A be an R-matrix whose determinant δ is a unit. Then we can find its inverse by Cramer's Rule: $\delta I = A(\operatorname{adj} A)$, where the adjoint matrix is calculated from A by taking determinants of minors [Chapter 1 (5.4)]. This rule also holds in any ring. So if δ is a unit, we can solve for A^{-1} in R as

$$A^{-1} = \delta^{-1}(\operatorname{adj} A).$$

(2.2) **Corollary.** The invertible $n \times n$ matrices A with entries in R are those matrices whose determinant is a unit. They form a group

$$GL_n(R) = \{\text{invertible } n \times n \ R\text{-matrices}\},$$

called the *general linear group over R*. □

The fact that the determinant of an invertible matrix must be a unit is a strong condition on the matrix when R has few units. For instance, if R is the ring of integers, the determinant must be ± 1. Most integer matrices are invertible real matrices, so they are in $GL_n(\mathbb{R})$. But unless the determinant ± 1, the entries of the inverse matrix won't be integers, so the inverses will not be in $GL_n(\mathbb{Z})$. Nevertheless, there are always reasonably many invertible matrices if $n > 1$, because the elementary matrices

$$I + ae_{ij} = \begin{bmatrix} 1 & & a \\ & \ddots & \\ & & 1 \end{bmatrix}, \quad i \neq j, \quad a \in R,$$

have determinant 1. These matrices generate a good-sized group. The other elementary matrices, the transposition matrices and the matrices

$$\begin{bmatrix} 1 & & \\ & \ddots u \ddots & \\ & & 1 \end{bmatrix}, \quad u \quad \text{a unit in } R,$$

are also invertible.

We now return to the discussion of modules over a ring R. The concepts of basis and independence (Chapter 3, Section 3) can be carried over from vector spaces to modules without change: An ordered set $(v_1, \dots, v_k)$ of elements of a module V is said to *generate* (or *span*) V if every $v \in V$ is a linear combination:

$$v = r_1 v_1 + \cdots + r_k v_k, \quad \text{with } r_i \in R. \tag{2.3}$$

In that case the elements v_i are called *generators*. A module V is said to be *finitely generated* if there exists a finite set of generators. Most of the modules we study will be finitely generated. A $\mathbb{Z}$-module V is finitely generated if and only if it is a finitely generated abelian group in the sense of Chapter 6, Section 8.

We saw in Section 1 that modules needn't be isomorphic to any of the modules R^k. However, a given module may happen to be, and if so, it is called a *free module* too. Thus a finitely generated module V is free if there is an isomorphism

$$\varphi: R^n \xrightarrow{\sim} V.$$

For instance, lattices in $\mathbb{R}^2$ are free $\mathbb{Z}$-modules, whereas finite, nonzero abelian groups are not free. A free $\mathbb{Z}$-module is also called a *free abelian group*. Free modules form an important and natural class, and we will study them first. We will study general modules beginning in Section 5.

Following the definitions for vector spaces, we call a set of elements $(v_1, \dots, v_n)$ of a module V *independent* if no nontrivial linear combination is zero, that is, if the following condition holds:

$$\textit{If } r_1 v_1 + \cdots + r_n v_n = 0, \textit{ with } r_i \in R, \textit{ then } r_i = 0 \text{ for } i = 1, \dots, n. \tag{2.4}$$

The set is a *basis* if it is both independent and a generating set. The *standard basis* $\mathbf{E} = (e_1, \dots, e_k)$ is a basis of R^k. Exactly as with vector spaces, $(v_1, \dots, v_k)$ is a basis if every $v \in V$ is a linear combination (2.3) in a unique way.

We may also speak of linear combinations and linear independence of infinite sets, using the terminology of Chapter 3, Section 5.

Let us denote the ordered set $(v_1, \dots, v_n)$ by $\mathbf{B}$, as in Chapter 3, Section 3. Then multiplication by $\mathbf{B}$,

$$\mathbf{B}X = (v_1, \dots, v_n) \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = v_1 x_1 + \cdots + v_n x_n,$$

defines a homomorphism of modules

$$\mu: R^n \longrightarrow V. \tag{2.5}$$

This homomorphism is surjective if and only if the set $(v_1, \dots, v_n)$ generates V, and injective if and only if it is independent. Thus it is bijective if and only if $\mathbf{B}$ is a basis of V, in which case V is a free module. So a module V has a basis if and only if it is free. Most modules have no bases.

Computation with bases of free R-modules can be done in much the same way as with bases of vector spaces, using matrices with entries in R. In particular, we can speak of the *coordinate vector* of an element $v \in V$, with respect to a basis $\mathbf{B} = (v_1, \ldots, v_n)$. It is the unique column vector $X \in R^n$ such that

$$v = \mathbf{B}X = v_1x_1 + \cdots + v_nx_n.$$

If two bases $\mathbf{B} = (v_1, \ldots, v_n)$ and $\mathbf{B}' = (v_1', \ldots, v_r')$ for the same free module V are given, then the matrix of change of basis is obtained as in Chapter 3, Section 4 by writing the elements v_j of the first basis as linear combinations of the second basis: $\mathbf{B} = \mathbf{B}'P$, or

$$(2.6) \qquad v_j = \sum_{i=1}^{r} v_i' p_{ij}.$$

As with vector spaces, any two bases of the same free module over a nonzero ring have the same cardinality, provided that R is not the zero ring. Thus $n = r$ in the above bases. This can be proved by considering the inverse matrix $Q = (q_{ij})$ which is obtained by writing $\mathbf{B}'$ in terms of $\mathbf{B}$: $\mathbf{B}' = \mathbf{B}Q$. Then

$$\mathbf{B} = \mathbf{B}'P = \mathbf{B}QP.$$

Since $\mathbf{B}$ is a basis, there is only one way to write v_j as a linear combination of $(v_1, \ldots, v_n)$, and that is $v_j = 1v_j$, or $\mathbf{B} = \mathbf{B}I$. Therefore $QP = I$, and similarly $PQ = I$: The matrix of change of basis is an invertible R-matrix.

Now P is an $r \times n$ matrix, and Q is a $n \times r$ matrix. Suppose that $r > n$. Then we make P and Q square by adding zeros:

$$\left[\begin{array}{c|c} P & 0 \end{array}\right]\left[\begin{array}{c} Q \\ \hline 0 \end{array}\right] = I.$$

This does not change the product PQ. But the determinants of these square matrices are zero, so they are not invertible, because $R \neq 0$. This shows that $r = n$, as claimed.

It is a startling fact that there exist *noncommutative* rings R for which the modules R^n for $n = 1, 2, 3, \ldots$ are all isomorphic (see miscellaneous exercise 6). Determinants do not work well unless the matrix entries commute.

Unfortunately, most concepts relating to vector spaces have different names when used for modules over rings, and it is too late to change them. The number of elements of a basis for a free module V is called the *rank* of V, instead of the dimension.

As we have already remarked, every homomorphism $\varphi: R^n \longrightarrow R^m$ between column vectors is left multiplication by a matrix A. If $\varphi: V \longrightarrow W$ is a homomorphism of free R-modules with bases $\mathbf{B} = (v_1, \ldots, v_n)$ and $\mathbf{C} = (w_1, \ldots, w_m)$ respectively, then the *matrix* of the homomorphism is defined to be $A = (a_{ij})$, where

$$(2.7) \qquad \varphi(v_j) = \sum_i w_i a_{ij}$$

as before [Chapter 4 (2.3)]. A change of the bases **B**, **C** by invertible R-matrices P, Q changes the matrix of φ to $A' = QAP^{-1}$ [Chapter 4 (2.7)].

3. THE PRINCIPLE OF PERMANENCE OF IDENTITIES

In this section, we address the following question: Why do the properties of matrices with entries in a field continue to hold when the entries are in an arbitrary ring? Briefly, the reason is that they are *identities*, which means that they hold when the matrix entries are replaced by variables. To be more precise, assume we want to prove some identity such as the multiplicative property of the determinant, $(\det A)(\det B) = \det(AB)$, or Cramer's Rule. Suppose that we have already checked the identity for matrices with complex entries. We don't want to do the work again, and anyhow we may have used special properties of $\mathbb{C}$, such as the field axioms, the fact that every complex polynomial has a root, or the fact that $\mathbb{C}$ has characteristic zero, to check the identity there. We did use special properties to prove the identities mentioned, so the proofs we gave will not work for rings. We are now going to show how to deduce such identities for all rings from the same identities for the complex numbers.

The principle is very general, but in order to focus attention, let us concentrate on the identity $(\det A)(\det B) = \det(AB)$. We begin by replacing the matrix entries with variables. So we consider the same identity

$$(\det X)(\det Y) = \det(XY),$$

where X and Y denote $n \times n$ matrices with variable entries. Then we can substitute elements in any ring R for these variables. Formally, the substitution is defined in terms of the ring of integer polynomials $\mathbb{Z}[\{x_{ij}\}, \{y_{k\ell}\}]$ in $2n^2$ variable matrix entries. There is a unique homomorphism from the ring of integers to any ring R [Chapter 10 (3.9)]. Given matrices $A = (a_{ij})$, $B = (b_{k\ell})$ with entries in R, there is a homomorphism

$$\mathbb{Z}[\{x_{ij}\}, \{y_{k\ell}\}] \longrightarrow R, \tag{3.1}$$

the substitution homomorphism, which sends $x_{ij} \rightsquigarrow a_{ij}$ and $y_{k\ell} \rightsquigarrow b_{k\ell}$ [Chapter 10 (3.4)]. Our variable matrices have entries in the polynomial ring, and it is natural to say that the homomorphism sends $X \rightsquigarrow A$ and $Y \rightsquigarrow B$, meaning that the entries of $X = (x_{ij})$ are mapped to the entries of $A = (a_{ij})$ and so on, by the map.

The general principle we have in mind is this: Suppose we want to prove an identity, all of whose terms are polynomials with integer coefficients in the matrix entries. Then the terms are compatible with ring homomorphisms: For example, if a homomorphism $\varphi: R \longrightarrow R'$ sends $A \rightsquigarrow A'$ and $B \rightsquigarrow B'$, then it sends $\det A \rightsquigarrow \det A'$. To see this, note that the complete expansion of the determinant is

$$\det A = \sum_p \pm a_{1p(1)} \cdots a_{np(n)},$$

the summation being over all permutations p. Since φ is a homomorphism,

$$\varphi(\det A) = \sum_p \pm \varphi(a_{1p(1)} \cdots a_{np(n)}) = \sum \pm a_{1p(1)}' \cdots a_{np(n)}' = \det A'.$$

Obviously, this is a general principle. Consequently, if our identity holds for the R-matrices A, B, then it also holds for the R'-matrices A', B'.

Now for every pair of matrices A, B, we have the homomorphism (3.1) which sends $X \rightsquigarrow A$ and $Y \rightsquigarrow B$. We substitute $\mathbb{Z}[\{x_{ij}\}, \{y_{ij}\}]$ for R and R for R' in the principle just described. We conclude that if the identity holds for the variable matrices X, Y in $\mathbb{Z}[\{x_{ij}\}, \{y_{ij}\}]$, then it holds for every pair of matrices in any ring R:

(3.2) *To prove our identity in general, we need only prove it for the variable matrices X, Y in the ring $\mathbb{Z}[\{x_{ij}\}, \{y_{ij}\}]$.*

To prove it for variable matrices, we consider the ring of integers as a subring of the field of complex numbers, noting the inclusion of polynomial rings

$$\mathbb{Z}[\{x_{ij}\}, \{y_{ij}\}] \subset \mathbb{C}[\{x_{ij}\}, \{y_{ij}\}].$$

We may as well check our identity in the bigger ring. Now by hypothesis, our identity is equivalent to the equality of certain polynomials in the variables $\{x_{ij}\}, \{y_{ij}\}, \ldots$. Let us write the identity as $f(x_{ij}, y_{kl}) = 0$. The symbol f may stand for several polynomials.

We now consider the polynomial *function* corresponding to the polynomial $f(x_{ij}, y_{kl})$, call it $\tilde{f}(x_{ij}, y_{kl})$. If the identity has been proved for all complex matrices, then it follows that $\tilde{f}(x_{ij}, y_{kl})$ is the zero function. We apply the fact [Chapter 10 (3.8)] that a polynomial is determined by the function it defines to conclude that $f(x_{ij}, y_{ij}) = 0$, and we are done.

It is possible to formalize the above discussion and to prove a precise theorem concerning the validity of identities in an arbitrary ring. However, even mathematicians occasionally feel that it isn't worthwhile making a precise formulation—that it is easier to consider each case as it comes along. This is one of those occasions.

4. DIAGONALIZATION OF INTEGER MATRICES

In this section we discuss simplification of an $m \times n$ integer matrix $A = (a_{ij})$ by a succession of elementary operations. We will apply this procedure later to classify abelian groups. The same method will work for matrices with entries in a Euclidean domain and, with some modification, for matrices with entries in a principal ideal domain.

The best results are obtained if we allow both row and column operations together. So we allow these operations:

(4.1)

(i) add an integer multiple of one row to another, or add an integer multiple of one column to another;
(ii) interchange two rows or two columns;
(iii) multiply a row or a column by a unit.

Of course, the units in $\mathbb{Z}$ are ± 1. Any such operation can be made by multiplying A on the left or right by a suitable elementary integer matrix. The result of a sequence of these operations will have the form

$$A' = QAP^{-1}, \tag{4.2}$$

where $Q \in GL_m(\mathbb{Z})$ and $P^{-1} \in GL_n(\mathbb{Z})$ are products of elementary integer matrices. Needless to say, we could drop the inverse symbol from P. We put it there because we will want to interpret the operation as a change of basis.

Over a *field*, any matrix can be brought into the block form

$$A' = \begin{bmatrix} I & \\ & 0 \end{bmatrix}$$

by such operations [Chapter 4 (2.9)]. We can not hope for such a result when working with integers. We can't even do it for 1×1 matrices. But we can diagonalize:

(4.3) **Theorem.** Let A be an $m \times n$ integer matrix. There exist products Q, P of elementary integer matrices as above, so that $A' = QAP^{-1}$ is diagonal:

$$\begin{bmatrix} \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_r \end{bmatrix} & \\ & 0 \end{bmatrix}$$

where the diagonal entries d_i are nonnegative and where each diagonal entry divides the next: $d_1 \mid d_2,\ d_2 \mid d_3, \ldots$.

Proof. The strategy is to perform a sequence of operations so as to end up with a matrix

$$\begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{bmatrix} \tag{4.4}$$

in which d_1 divides every entry of B. When this is done, we work on B. The process is based on repeated division with remainder. We will describe a systematic method, though using this method is usually not the quickest way to proceed.

We may assume $A \neq 0$.

Step 1: By permuting rows and columns, move a nonzero entry with smallest absolute value to the upper left corner. Multiply the first row by -1 if necessary, so that this upper left entry a_{11} becomes positive.

We now try to clear out the first row and column. Whenever an operation produces a nonzero entry in the matrix whose absolute value is smaller than $|a_{11}|$, we go back to Step 1 and start the whole process over. This is likely to spoil the work we have done to clear out matrix entries. However, progress is being made because the size of a_{11} is reduced every time. We will not have to return to Step 1 infinitely often.

Step 2: Choose a nonzero entry a_{i1} in the first column, with $i > 1$, and divide by a_{11}:

$$a_{i1} = a_{11}q + r,$$

where $0 \leq r < a_{11}$. Subtract q times (row 1) from (row i). This changes a_{i1} to r.

If $r \neq 0$, we go back to Step 1. If $r = 0$, we have produced a zero in the first column. Finitely many repetitions of Steps 1 and 2 result in a matrix in which $a_{i1} = 0$ for all $i > 1$. Similarly, we may use the analogue of Step 2 for column operations to clear out the first row, eventually ending up with a matrix in which the only nonzero entry in the first row and column is a_{11}, as required by (4.3). However, a_{11} may not yet divide every entry of the matrix B (4.4).

Step 3: Assume that a_{11} is the only nonzero entry in the first row and column, but that some entry b of B is not divisible by a_{11}. Add the column of A which contains b to column 1. This produces an entry b in the first column.

We go back to Step 2. Division with remainder will now produce a smaller matrix entry, sending us back to Step 1. A finite sequence of these steps will produce a matrix of the form (4.4), allowing us to proceed by induction. □

(4.5) **Example.** We do not follow the systematic method:

$$A = \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \xrightarrow[\text{oper}]{\text{column}} \begin{bmatrix} 1 & -1 \\ 3 & 2 \end{bmatrix} \xrightarrow[\text{oper}]{\text{column}} \begin{bmatrix} 1 & \\ 3 & 5 \end{bmatrix} \xrightarrow[\text{oper}]{\text{row}} \begin{bmatrix} 1 & \\ & 5 \end{bmatrix} = A'.$$

Here

$$Q = \begin{bmatrix} 1 & \\ -3 & 1 \end{bmatrix} \quad \text{and} \quad P^{-1} = \begin{bmatrix} 1 & \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

Note that the key ingredient in this proof is the division algorithm. The same proof will work when $\mathbb{Z}$ is replaced by any Euclidean domain.

(4.6) **Theorem.** Let R be a Euclidean domain, for instance a polynomial ring $F[t]$ in one variable over a field. Let A be an $m \times n$ matrix with entries in R. There are products Q, P of elementary R-matrices such that $A' = QAP^{-1}$ is diagonal and such

that each diagonal entry of A' divides the next: $d_1 | d_2 | d_3 | \dots$. If $R = F[t]$, we can normalize by requiring the polynomials d_i to be monic. □

(4.7) **Example.** Diagonalization of a matrix of polynomials:

$$A = \begin{bmatrix} t^2-3t+2 & t-2 \\ (t-1)^3 & t^2-3t+2 \end{bmatrix} \xrightarrow[\text{oper}]{\text{row}} \begin{bmatrix} t^2-3t+2 & t-2 \\ (t-1)^2 & 0 \end{bmatrix} \xrightarrow[\text{oper}]{\text{row}} \begin{bmatrix} -t+1 & t-2 \\ (t-1)^2 & 0 \end{bmatrix} \xrightarrow[\text{oper}]{\text{column}}$$

$$\begin{bmatrix} -1 & t-2 \\ (t-1)^2 & 0 \end{bmatrix} \xrightarrow[\text{oper}]{\text{column}} \begin{bmatrix} -1 & 0 \\ (t-1)^2 & (t-1)^2(t-2) \end{bmatrix} \xrightarrow[\text{oper}]{\text{row}} \begin{bmatrix} 1 & \\ & (t-1)^2(t-2) \end{bmatrix} = A'.$$

In both examples, we ended up with 1 in the upper left corner. This isn't surprising. The matrix entries will often have greatest common divisor 1.

The diagonalization of integer matrices can be used to describe homomorphisms between free abelian groups. As we have already remarked (2.8), a homo*morphism $\varphi: V \longrightarrow W$ of free abelian groups is described by a matrix*, once bases for V and W are chosen. A change of bases in V, W by invertible integer matrices P, Q changes A to $A' = QAP^{-1}$. So we have proved the following theorem:

(4.8) **Theorem.** Let $\varphi: V \longrightarrow W$ be a homomorphism of free abelian groups. There exist bases of V and W such that the matrix of the homomorphism has the diagonal form (4.3). □

In the rest of this section, we will investigate the meaning of this theorem for two auxiliary groups associated to a homomorphism: its kernel and its image.

Let $\varphi: \mathbb{Z}^n \longrightarrow \mathbb{Z}^m$ be left multiplication by the $m \times n$ integer matrix A. The kernel of φ is the subgroup of $\mathbb{Z}^n$ of integer solutions of the system of linear equations

$$AX = 0. \tag{4.9}$$

These solutions can be read off immediately when the matrix is diagonal: In order for X to solve the diagonal system $d_1x_1 = 0, \dots, d_nx_n = 0$, we must have $x_i = 0$ unless $d_i = 0$, and if $d_i = 0$, then x_i can be arbitrary.

To solve (4.9) in general, we may diagonalize A, say to $A' = QAP^{-1}$, where Q, P are products of elementary integer matrices. We make the change of variable $X' = PX$ and solve the diagonal system

$$A'X' = QAP^{-1}X' = 0.$$

Since Q is invertible, the system of equations $QAX = 0$ has the same solutions as the system $AX = 0$. So the solutions of the original system are $X = P^{-1}X'$.

Next, let us examine the image of $\varphi: \mathbb{Z}^n \longrightarrow \mathbb{Z}^m$, the map defined by multiplication by the integer matrix A as before. We can describe this image as the set of vectors $B \in \mathbb{Z}^m$ such that the system of integer equations $AX = B$ has an integer solution. We will often denote this image by $A\mathbb{Z}^n$. Multiplication by A sends the basis

vectors $e_1, \ldots, e_n \in \mathbb{Z}^n$ to the columns

$$(4.10) \qquad A_1 = \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}, \ldots, A_n = \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

of A, so the image is the set of integer linear combinations of these columns. In other words, the columns generate the image.

We can turn this description around, starting with an arbitrary *subgroup* S of the free abelian group $\mathbb{Z}^m$ which is given to us explicitly by a set of generators $A_1, \ldots, A_n \in \mathbb{Z}^m$. Let A be the matrix whose columns are A_i. Then S is the image of left multiplication by A. This interpretation of S as the image of a homomorphism tells us the meaning of left and right multiplication by invertible integer matrices Q and P^{-1}: Left multiplication by Q corresponds to a change of basis in the module $\mathbb{Z}^m$, the range of the map. Its effect is to multiply each of the generators A_i by Q. On the other hand, right multiplication by P^{-1} represents a change of basis in the domain $\mathbb{Z}^n$. This changes the generating set of S. For example, adding r times column 1 to column 2 changes A_2 to $A_2' = A_2 + rA_1$ and leaves the other generators unchanged. Combining these observations with diagonalization results in the following theorem:

(4.11) **Theorem.** Let S be a subgroup of a free abelian group W of rank m. There is a basis $(w_1, \ldots, w_m)$ of W and a basis $(u_1, \ldots, u_n)$ of S with the following properties: (i) $n \leq m$, (ii) for each $j \leq n$ there is a positive integer d_j such that $u_j = d_j w_j$, and (iii) $d_1 \mid d_2 \mid d_3 \ldots$.

(4.12) **Corollary.** Every subgroup of a free abelian group of rank m is free, and its rank is at most m. □

Proof of Theorem (4.11). Roughly speaking, we need only choose a basis $\mathbf{B} = (w_1, \ldots, w_m)$ for W and a set of generators $(u_1, \ldots u_n)$ for S, to obtain an $m \times n$ matrix A which represents S as above. The diagonalization theorem gives us a diagonal matrix $A' = QAP^{-1}$ representing S with respect to a new basis $\mathbf{B}' = (w_1', \ldots, w_p')$ and new generating set $(u_1', \ldots, u_n')$. Then $u_j' = d_j w_j'$. We drop the primes to obtain the basis and generating set required. This completes the proof except for three points.

First, we may have $n > m$, that is, there may be more columns than rows. But if so, then since A' is diagonal, its jth column is zero for each $j > m$; hence the corresponding generator u_j is zero too. The zero element is useless as a generator, so we throw it out. For the same reason, we may throw out a generator u_j whenever $d_j = 0$. After we do this, all d_j will be positive, and we will have $n \leq m$.

Notice that if S is the zero subgroup, we will end up throwing out all the generators. As with vector spaces, we must adopt the convention that the empty set generates the zero module, or else make a special mention of this exceptional case in the statement of the theorem.

Next, we verify that if the basis and generating set are chosen so that $d_i > 0$ and $n \leq m$, then $(u_1, \ldots, u_n)$ is a basis of S. Since it generates S, what has to be proved is that $(u_1, \ldots, u_n)$ is independent. We rewrite a linear relation $r_1 u_1 + \cdots + r_n u_n = 0$ in the form $r_1 d_1 w_1 + \cdots + r_n d_n w_n = 0$. Since $(w_1, \ldots, w_m)$ is a basis, $r_i d_i = 0$ for each i, and since $d_i > 0$, $r_i = 0$.

The final point is more serious: We need a finite set of generators of S to get started. How do we know that there is such a set? It is a fact that every subgroup of a finitely generated abelian group is itself finitely generated. We will prove this in Section 5. For the moment, the theorem is proved only with the additional hypothesis that S is finitely generated. The hypothesis that W is finitely generated can not be removed. □

Theorem (4.11) is quite explicit. Let S be the subgroup of $\mathbb{Z}^m$ generated by the columns of a matrix A, and suppose that $A' = QAP^{-1}$ is diagonal. To display S in the form asserted in the theorem, we rewrite this equation in the form

$$Q^{-1}A' = AP^{-1}, \tag{4.13}$$

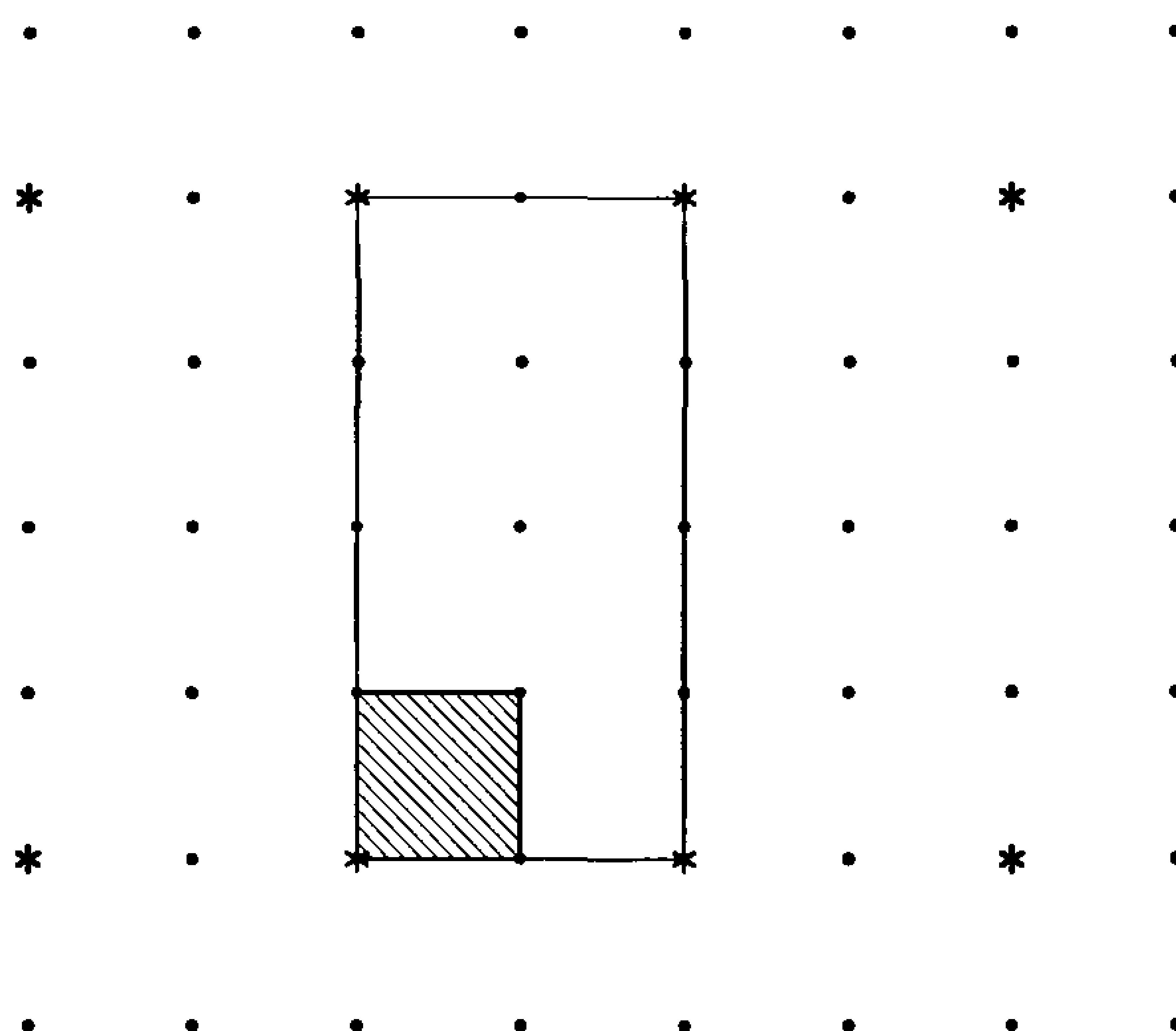
and we interpret it as follows: The columns of the matrix AP^{-1} form our new set of generators for S. Since the matrix A' is diagonal, (4.13) tells us that the new generators are multiples of the columns of Q^{-1}. We change the basis of $\mathbb{Z}^m$ from the standard basis to the basis made up of the columns of Q^{-1}. The matrix of this change of basis is Q [see Chapter 3 (4.21)]. Then the new generators are multiples of the new basis elements.

For instance, let S be the lattice in $\mathbb{R}^2$ generated by the two columns of the matrix A of Example (4.5): Then

$$Q^{-1}A' = \begin{bmatrix} 1 & \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1 & \\ & 5 \end{bmatrix} = \begin{bmatrix} 1 & \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = AP^{-1}. \tag{4.14}$$

The new basis of $\mathbb{Z}^2$ is $(w_1', w_2') = \left(\begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} \\ 1 \end{bmatrix}\right)$, and the new generators of S are $(u_1', u_2') = (u_1, u_2)P^{-1} = (w_1', 5w_2')$.

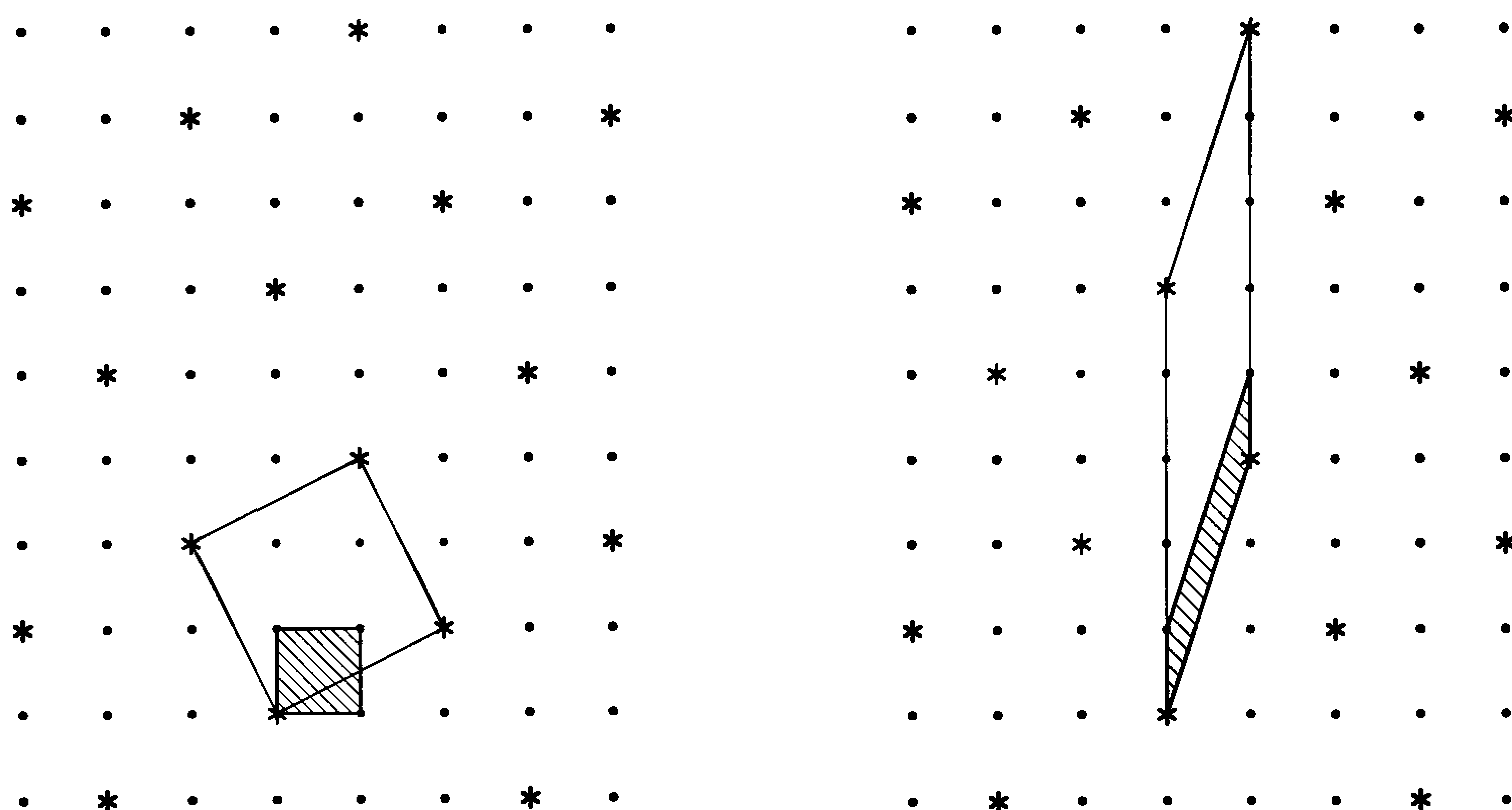
Theorem (4.3) is striking when it is used to describe the relative position of a sublattice S in a lattice L. To illustrate this, it will be enough to consider plane lattices. The theorem asserts that there are bases (v_1, v_2) and (w_1, w_2) of L and S such that the coordinate vectors of w_j with respect to the basis (v_1, v_2) are diagonal. Let us refer the lattice L back to $\mathbb{Z}^2 \subset \mathbb{R}^2$ by means of the basis (v_1, v_2). Then the equations $w_i = d_i v_i$ show that S looks like this figure, in which we have taken $d_1 = 2$ and $d_2 = 4$:



(4.15) **Figure.** $S = *$, matrix $\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$.

Notice the fact, which we have asserted before [Chapter 11 (10.10)], that the index $[L:S]$ is the ratio of the areas of the parallelograms spanned by bases. This is evident when the bases are in such a relative position.

In practice, when the lattices L and S are given to us in $\mathbb{R}^2$ at the start, the change of basis required to get such "commensurable" bases of L and S leads to rather long and thin parallelograms, as is shown below for Example (4.14).



(4.16) **Figure.** Diagonalization, applied to a sublattice.

5. GENERATORS AND RELATIONS FOR MODULES

In this section we turn our attention to modules which are not free. We will show how to describe a large class of modules by means of matrices called *presentation matrices*. We will then apply the diagonalization procedure to these matrices to the study of abelian groups.

As an example to keep in mind, we may consider an abelian group or $\mathbb{Z}$-module V which is generated by three elements (v_1, v_2, v_3). We suppose that these generators are subject to the relations

$$\begin{aligned} 3v_1 + 2v_2 + v_3 &= 0 \\ 8v_1 + 4v_2 + 2v_3 &= 0 \\ 7v_1 + 6v_2 + 2v_3 &= 0 \\ 9v_1 + 6v_2 + v_3 &= 0. \end{aligned} \tag{5.1}$$

The information describing this module is summed up in the matrix

$$A = \begin{bmatrix} 3 & 8 & 7 & 9 \\ 2 & 4 & 6 & 6 \\ 1 & 2 & 2 & 1 \end{bmatrix}, \tag{5.2}$$

whose columns are the coefficients of the relations (5.1):

$$(v_1, v_2, v_3)A = (0,0,0,0).$$

As usual, scalars appear on the right side in this matrix product. It is this method of describing a module which we plan to formalize.

If $(v_1, ,\ldots, v_m)$ are elements of an R-module V, equations of the form

$$a_1v_1 + \cdots + a_mv_m = 0, \quad a_i \in R, \tag{5.3}$$

are called *relations* among the elements. Of course, when we refer to (5.3) as a relation, we mean that the formal expression is a relation: If we evaluate it in V, we get $0 = 0$. Since the relation is determined by the R-vector $(a_1,\ldots, a_m)^t$, we will refer to this vector as a *relation vector*, meaning that (5.3) is true in V. By a *complete set of relations* we mean a set of relation vectors such that every relation vector is a linear combination of this set. It is clear that a matrix such as (5.2) will not describe the module V completely, unless its columns form a complete set of relations.

The concept of a complete set of relations can be confusing. It becomes much clearer when we work with homomorphisms of free modules rather than directly with the relations or the relation vectors. Let an $m \times n$ matrix A with entries in a ring R be given. As we know, left multiplication by this matrix is a homomorphism of R-modules

$$\varphi: R^n \longrightarrow R^m. \tag{5.4}$$

In addition to the kernel and image, which we described in the last section when $R = \mathbb{Z}$, there is another important auxiliary module associated with a homomorphism $\varphi: W \longrightarrow W'$ of R-modules, called its *cokernel*. The cokernel of φ is defined to be the quotient module

(5.5) $$W'/(\text{im } \varphi).$$

If we denote the image of left multiplication by A by AR^n, the cokernel of (5.4) is R^m/AR^n. This cokernel is said to be *presented* by the matrix A. More generally, we will call any isomorphism

(5.6) $$\sigma: R^m/AR^n \xrightarrow{\sim} V$$

a *presentation* of a module V, and we say that the matrix A is a *presentation matrix* for V if there is such an isomorphism.

For example, the cyclic group $\mathbb{Z}/(5)$ is presented as a $\mathbb{Z}$-module by the 1×1 integer matrix $[5]$. As another example, let V be the $\mathbb{Z}$-module presented by the matrix $\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}$. The columns of this matrix are the relation vectors, so V is generated by two elements v_1, v_2 with the relations $2v_1 + v_2 = -v_1 + 2v_2 = 0$. We may solve the first relation, obtaining $v_2 = -2v_1$. This allows us to eliminate the second generator. Substitution into the second relation gives $-5v_1 = 0$. So V can also be generated by a single generator v_1, with the single relation $5v_1 = 0$. This shows that V is isomorphic to $\mathbb{Z}/(5)$. This 2×2 matrix also presents the cyclic group $\mathbb{Z}/(5)$.

We will now describe a theoretical method of finding a presentation of a given module V. To carry out this method in practice, the module would have to be given in a very explicit way. Our first step is to choose a set of generators $(v_1, \dots, v_m)$. So V must be finitely generated for us to get started. These generators provide us with a surjective homomorphism

(5.7) $$f: R^m \longrightarrow V,$$

sending the column vector $X = (x_1, \dots, x_m)$ to $v_1x_1 + \cdots + v_mx_m$. The elements of the kernel of f are the relation vectors. Let us denote this kernel by W. By the First Isomorphism Theorem, V is isomorphic to R^m/W.

We repeat the procedure, choosing a set of generators $(w_1, \dots, w_n)$ for W, and we use these generators to define a surjective homomorphism

(5.8) $$g: R^n \longrightarrow W$$

as before. Since W is a submodule of R^m, composition of the homomorphism g with the inclusion $W \subset R^m$ gives us a homomorphism

(5.9) $$\varphi: R^n \longrightarrow R^m.$$

This homomorphism is left multiplication by a matrix A. By construction, W is the image of φ, which is AR^n, so $R^m/AR^n = R^m/W \approx V$. Therefore, A is a presentation matrix for V.

The columns of the matrix A are our chosen generators for the module W of relations:

$$w_1 = \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}, \ldots, w_n = \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

Since they generate W, these columns form a complete set of relations among the generators $(v_1, \ldots, v_m)$ of the module V. Since the columns are relation vectors,

(5.10) $$(v_1, \ldots, v_m)A = 0.$$

Thus the presentation matrix A for a module V is determined by

(5.11)

(i) a set of generators for V, and
(ii) a complete set of relations among these generators.

We have let one point slip by in this description. In order to have a finite set of generators for the module of relations W, this module must be finitely generated. This does not look like a satisfactory hypothesis, because the relationship of our original module V with W is unclear. We don't mind assuming that V is finitely generated, but it isn't good to impose hypotheses on a module which arises in the course of some auxiliary construction. We will need to examine this point more closely [see (5.16)]. But except for this point, we can now speak of generators and relations for a finitely generated R-module V.

Since the presentation matrix depends on the choices (5.11), many matrices present the same module, or isomorphic modules. Here are some rules for manipulating a matrix A without changing the isomorphism class of the module it presents:

(5.12) **Proposition.** Let A be an $m \times n$ presentation matrix for a module V. The following matrices A' present the same module V:

(i) $A' = QAP^{-1}$, where $Q \in GL_m(R)$ and $P \in GL_n(R)$;
(ii) A' is obtained by deleting a column of zeros;
(iii) the jth column of A is e_i, and A' is obtained from A by deleting the ith row and jth column.

Proof.

(i) The module R^m/AR^n is isomorphic to V. Since the change of A to QAP^{-1} corresponds to a change of basis in R^m and R^n, the isomorphism class of the quotient module does not change.
(ii) A column of zeros corresponds to the trivial relation, which can be omitted.
(iii) Suppose that the jth column of the matrix A is e_i. The corresponding relation is $v_i = 0$. So it holds in the module V, and therefore v_i can be left out of the gen-

erating set $(v_1, \ldots, v_m)$. Doing so changes the matrix A by deleting the ith row and jth column. □

It may be possible to simplify a matrix quite a lot by these rules. For instance, our original example of the integer matrix (5.2) reduces as follows:

$$A = \begin{bmatrix} 3 & 8 & 7 & 9 \\ 2 & 4 & 6 & 6 \\ 1 & 2 & 2 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 & 2 & 1 & 6 \\ 0 & 0 & 2 & 4 \\ 1 & 2 & 2 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 2 & 1 & 6 \\ 0 & 2 & 4 \end{bmatrix} \longrightarrow \begin{bmatrix} 2 & 1 & 6 \\ -4 & 0 & -8 \end{bmatrix} \longrightarrow$$

$$\longrightarrow [-4 \quad -8] \longrightarrow [-4 \quad 0] \longrightarrow [4].$$

Thus A presents the abelian group $\mathbb{Z}/(4)$.

By definition, an $m \times n$ matrix presents a module by means of m generators and n relations. But as we see from this example, the number of generators and the number of relations depend on choices. They are not uniquely determined by the module.

Consider two more examples: The 2×1 matrix $\begin{bmatrix} 4 \\ 0 \end{bmatrix}$ presents an abelian group V by means of two generators (v_1, v_2) and one relation $4v_1 = 0$. We can not simplify this matrix. The group which it presents is isomorphic to the product group $\mathbb{Z}/(4) \times \mathbb{Z}$. On the other hand, the matrix $[4 \quad 0]$ presents a group with one generator v_1 and two relations, the second of which is the trivial relation. This group is $\mathbb{Z}/(4)$.

We will now discuss the problem of finite generation of the module of relations. For modules over a nasty ring, this module needn't be finitely generated, even though V is. Fortunately this problem does not occur with the rings we have been studying, as we will now show.

(5.13) **Proposition.** The following conditions on an R-module V are equivalent:

(i) Every submodule W of V is finitely generated;

(ii) *ascending chain condition:* There is no infinite strictly increasing chain $W_1 < W_2 < \ldots$ of submodules of V.

Proof. Assume that V satisfies the ascending chain condition, and let W be a submodule of V. We select a set $w_1, w_2, \ldots, w_k$ of generators of W in the following way: If $W = 0$, then W is generated by the empty set. If not, we start with a nonzero element $w_1 \in W$. To continue, assume that $w_1, \ldots, w_i$ have been chosen, and let W_i be the submodule generated by these elements. If W_i is a proper submodule of W, let w_{i+1} be an element of W which is not contained in W_i. Then $W_1 < W_2 < \ldots$. Since V satisfies the ascending chain condition, this chain of submodules can not be continued indefinitely. Therefore some W_k is equal to W. Then $(w_1, \ldots, w_k)$ generates W. The converse follows the proof of Theorem (2.10) of Chapter 11. Assume that every

submodule of V is finitely generated, and let $W_1 \subset W_2 \subset \ldots$ be an infinite increasing chain of submodules of V. Let U denote the union of these submodules. Then U is a submodule [see Chapter 11 (2.11)]; hence it is finitely generated. Let $u_1, \ldots u_r$ be generators for U. Each u_ν is in one of the modules W_i, and since the chain is increasing, there is an i such that all of the generators are in W_i. Then the module U they generate is also in W_i, and we have $U \subset W_i \subset W_{i+1} \subset U$. This shows that $U = W_i = W_{i+1}$ and that the chain is not strictly increasing. □

(5.14) **Lemma.**

(a) Let $\varphi: V \longrightarrow W$ be a homomorphism of R-modules. If the kernel and the image of φ are finitely generated modules, so is V. If V is finitely generated and if φ is surjective, then W is finitely generated. More precisely, suppose that $(v_1, \ldots, v_n)$ generates V and that φ is surjective. Then $(\varphi(v_1), \ldots, \varphi(v_n))$ generates W.

(b) Let W be a submodule of an R-module V. If both W and V/W are finitely generated, so is V. If V is finitely generated, so is V/W.

Proof. For the first assertion of (a), we follow the proof of the dimension formula for linear transformations [Chapter 4 (1.5)], choosing a set of generators $(u_1, \ldots u_k)$ for ker φ and a set of generators $(w_1, \ldots, w_m)$ for im φ. We also choose elements $v_i \in V$ such that $\varphi(v_i) = w_i$. Then we claim that the set $(u_1, \ldots, u_k; v_1, \ldots, v_m)$ generates V. Let $v \in V$ be arbitrary. Then $\varphi(v)$ is a linear combination of $(w_1, \ldots, w_m)$, say $\varphi(v) = a_1 w_1 + \cdots + a_m w_m$. Let $v' = a_1 v_1 + \cdots + a_m v_m$. Then $\varphi(v') = \varphi(v)$. Hence $v - v' \in \ker \varphi$, so $v - v'$ is a linear combination of $(u_1, \ldots, u_k)$, say $v - v' = b_1 u_1 + \cdots + b_k u_k$. Therefore $v = a_1 v_1 + \cdots + a_m v_m + b_1 u_1 + \cdots + b_k u_k$. This shows that the set $(u_1, \ldots, u_k; v_1, \ldots, v_m)$ generates V, as required. The proof of the second assertion of (a) is easy. Part (b) follows from part (a) by a consideration of the canonical homomorphism $\pi: V \longrightarrow V/W$. □

(5.15) **Definition.** A ring R is called *noetherian* if every ideal of R is finitely generated.

Principal ideal domains are obviously noetherian, so the rings $\mathbb{Z}$, $\mathbb{Z}[i]$, and $F[x]$ (F a field) are noetherian.

(5.16) **Corollary.** Let R be a noetherian ring. Every proper ideal I of R is contained in a maximal ideal.

Proof. If I is not maximal itself, then it is properly contained in a proper ideal I_2, and if I_2 is not maximal, it is properly contained in a proper ideal I_3, and so on. By the ascending chain condition (5.13), the chain $I = I_1 < I_2 < I_3 \ldots$ must be finite. Therefore I_k is maximal for some k. □

The relevance of the notion of noetherian ring to our problem is shown by the following proposition:

(5.17) **Proposition.** Let V be a finitely generated module over a noetherian ring R. Then every submodule of V is finitely generated.

Proof. It suffices to prove the proposition in the case that $V = R^m$. For assume that we have proved that the submodules of R^m are finitely generated, for all m. Let V be a finitely generated R-module. Then there is a surjective map $\varphi: R^m \longrightarrow V$. Given a submodule S of V, let $L = \varphi^{-1}(S)$. Then L is a submodule of the module R^m, and hence L is finitely generated. Also, the map $L \longrightarrow S$ is surjective. Hence S is finitely generated (5.14).

To prove the proposition when $V = R^m$, we use induction on m. A submodule of R is the same as an ideal of R (1.3). Thus the noetherian hypothesis on R tells us that the proposition holds for $V = R^m$ when $m = 1$. Suppose $m > 1$. We consider the projection

$$\pi: R^m \longrightarrow R^{m-1}$$

given by dropping the last entry: $\pi(a_1,\dots,a_m) = (a_1,\dots,a_{m-1})$. Its kernel is $\{(0,\dots,0,a_m)\} \approx R$. Let $W \subset R^m$ be a submodule, and let $\varphi: W \longrightarrow R^{m-1}$ be the restriction of π to W. The image $\varphi(W)$ is finitely generated, by induction. Also, $\ker \varphi = (W \cap \ker \pi)$ is a submodule of $\ker \pi \approx R$, so it is finitely generated too. By Lemma (5.14), W is finitely generated, as required. □

This proposition completes the proof of Theorem (4.11).

Since principal ideal domains are noetherian, submodules of finitely generated modules over these rings are finitely generated. But in fact, most of the rings which we have been studying are noetherian. This follows from another of Hilbert's famous theorems:

(5.18) **Theorem.** *Hilbert Basis Theorem:* If a ring R is noetherian, then so is the polynomial ring $R[x]$.

The Hilbert Basis Theorem shows by induction that the polynomial ring $R[x_1,\dots,x_n]$ in several variables over a noetherian ring R is noetherian, hence that the rings $\mathbb{Z}[x_1,\dots,x_n]$ and $F[x_1,\dots,x_n]$ (F a field) are noetherian. Also, quotients of noetherian rings are noetherian:

(5.19) **Proposition.** Let R be a noetherian ring, and let I be an ideal of R. The quotient ring $\overline{R} = R/I$ is noetherian.

Proof. Let $\overline{J}$ be an ideal of $\overline{R}$, and let $J = \pi^{-1}(\overline{J})$ be the corresponding ideal of R, where $\pi: R \longrightarrow \overline{R}$ is the canonical map. Then J is finitely generated, say by $(a_1,\dots,a_m)$. It follows that the finite set $(\overline{a}_1,\dots,\overline{a}_m)$ generates $\overline{J}$ (5.14). □

Combining this proposition with the Hilbert Basis Theorem gives the following result:

(5.20) **Corollary.** Any ring which is a quotient of a polynomial ring over the integers or over a field is noetherian. □

Proof of the Hilbert Basis Theorem. Assume that R is noetherian, and let I be an ideal of the polynomial ring $R[x]$. We must show that a finite set of polynomials suffices to generate this ideal.

Let's warm up by reviewing the case that R is a field. In that case, we may choose a nonzero polynomial $f \in I$ of lowest degree, say

$$f(x) = a_n x^n + \cdots + a_1 x + a_0, \quad a_n \neq 0, \tag{5.21}$$

and prove that it generates the ideal as follows: Let

$$g(x) = b_m x^m + \cdots + b_1 x + b_0, \quad b_m \neq 0, \tag{5.22}$$

be a nonzero element of I. Then the degree m of g is at least n. We use induction on m. The polynomial

$$g(x) - (b_m/a_n)x^{m-n}f(x) = g_1(x) \tag{5.23}$$

is an element of I of degree $< m$. By induction, g_1 is divisible by f; hence g is divisible by f.

Formula (5.23) is the first step in the division with remainder of g by f. The method does not extend directly to arbitrary rings, because division with remainder requires that the leading coefficient of f be a unit. More precisely, in order to form the expression (5.23) we need to know that a_n divides b_m in the ring R, and there is no reason for this to be true. We will need more generators.

Let us denote by A the set of leading coefficients of all the polynomials in I, together with the zero element of R.

(5.24) **Lemma.** The set A of leading coefficients of the polynomials in an ideal of $R[x]$, together with 0, forms an ideal of R.

Proof. If $\alpha = a_n$ is the leading coefficient of f, then $r\alpha$ is the leading coefficient of rf, unless by chance $r\alpha = 0$. In both cases, $r\alpha \in A$. Next, let $\alpha = a_n$ be the leading coefficient of f, and let $\beta = b_m$ be the leading coefficient of g, where, say, $m \geq n$. Then α is also the leading coefficient of $x^{m-n}f$. Hence the coefficient of x^m in the polynomial $h = x^{m-n}f + g$ is $\alpha + \beta$. This is the leading coefficient of h unless it is zero, and in either case, $\alpha + \beta \in A$. □

We return to the proof of the Hilbert Basis Theorem. According to the lemma, the set A is an ideal of the noetherian ring R, so there exists a finite set of generators, say $(\alpha_1, \ldots, \alpha_k)$, for this ideal. We choose for each i, $1 \leq i \leq k$, a polynomial

$f_i \in I$ with leading coefficient α_i, and we multiply these polynomials by powers of x as necessary, so that their degrees become equal to some common integer n.

The set of polynomials $(f_1,\ldots,f_k)$ obtained in this way will allow us to adapt the induction step (5.23), but it will probably not generate I. We have little chance of finding a polynomial of degree $<n$ in the ideal $(f_1,\ldots,f_k)$. So we must add some elements of low degree to get generators for our ideal. The following lemma is easy, and we omit its proof:

(5.25) **Lemma.** Let P_n denote the set of polynomials in $R[x]$ which have degree $< n$, together with zero, and let $S_n = I \cap P_n$. Then S_n is an R-submodule of the R-module P_n.

The R-module P_n is generated by the monomials $1, x,\ldots, x^{n-1}$, so it is finitely generated. Since R is noetherian, we may use Lemma (5.25) and Proposition (5.17) to conclude that there is a finite set $(h_1,\ldots,h_s)$ of elements which generates S_n as an R-module. We claim that the combined set $(f_1,\ldots,f_k; h_1,\ldots,h_s)$ generates I.

Denote by J the ideal generated by this set. By construction, $J \subset I$. We need to prove the opposite inclusion, and we use induction on the degree of an element $g \in I$. We denote this degree by m. If $m < n$, then $g \in S_n$, and therefore g is a linear combination of $(h_1,\ldots,h_s)$, with coefficients in R. So $g \in J$ in that case. Assume that $m \geq n$, and let the leading coefficient of g be $b = b_m$. Then b is in the ideal A of leading coefficients, so it is a linear combination of the generators of that ideal, say $b = r_1\alpha_1 + \cdots + r_k\alpha_k$. Remembering that α_i is the leading coefficient of f_i, we see that the polynomial

$$p = x^{m-n}(\sum_i r_i f_i)$$

has the same leading coefficient and the same degree as g, and it is in J. So $g_1 = g - p$ has degree less than m. By induction, $g_1 \in J$, and hence $g \in J$. □

6. THE STRUCTURE THEOREM FOR ABELIAN GROUPS

The Structure Theorem for abelian groups asserts that a finitely generated abelian group V is a direct sum of cyclic groups. The work of the proof has already been done. We know that there exists a diagonal presentation matrix for V, and what remains for us to do is to interpret the meaning of this diagonal matrix for the group.

We first need to extend the concept of direct sum from vector spaces to arbitrary modules. The definition is the same. Let $W_1,\ldots, W_k$ be submodules of a module V. Their *sum* is the submodule which they generate. It consists of all sums

$$(6.1) \qquad W_1 + \cdots + W_k = \{v \in V \mid v = w_1 + \cdots + w_k, \quad \text{with } w_i \in W_i\}.$$

The verification that this is a submodule is routine, and it is the same as for sums of subspaces of a vector space. We say that V is the *direct sum* of the submodules W_i if

(6.2)

(i) they *generate*: $V = W_1 + \cdots + W_k$;

(ii) they are *independent*: If $w_1 + \cdots + w_k = 0$, with $w_i \in W_i$, then $w_i = 0$ for each i.

Thus V is the direct sum of the submodules W_i if every element $v \in V$ can be written uniquely in the form $v = w_1 + \cdots + w_k$, with $w_i \in W_i$. As with vector spaces, two submodules W_1, W_2 are independent if and only if $W_1 \cap W_2 = 0$ [see Chapter 3 (6.5)].

The symbol $\oplus$ is used to denote direct sums as before. So the notation

$$V = W_1 \oplus \cdots \oplus W_k \tag{6.3}$$

means that V is the direct sum of the submodules W_i.

(6.4) **Theorem.** *Structure Theorem for abelian groups:* Let V be a finitely generated abelian group. Then V is a direct sum of finite cyclic subgroups $C_{d_1}, \ldots, C_{d_k}$ and a free abelian group L:

$$V = C_{d_1} \oplus \cdots \oplus C_{d_k} \oplus L,$$

where the order d_i of C_{d_i} is greater than 1, and $d_1 \mid d_2 \mid d_3 \ldots$.

We will use additive notation for the law of composition in the cyclic group here. So C_n is generated by one element v, with one relation $nv = 0$. Thus C_n is isomorphic to $\mathbb{Z}/(n)$. The isomorphism $\mathbb{Z}/(n) \longrightarrow C_n$ sends the residue of an integer r to rv.

Proof of the theorem. We choose a presentation matrix A for V, determined by a set of generators and a complete set of relations. We can do this because V is finitely generated and because $\mathbb{Z}$ is a noetherian ring (see Section 5). By Proposition (5.12), the matrix A may be replaced by QAP^{-1}, where Q and P are invertible. Therefore we may assume that A is diagonal, that the diagonal entries are nonzero, and that each diagonal entry divides the next. Moreover, we can drop any column of zeros, and any row and column in which the diagonal entry is 1 (5.12). So we may assume that the diagonal entries d_i are not 0 or 1. The matrix A will then have the shape

$$\begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_k \\ \hline & & 0 & \end{bmatrix}. \tag{6.5}$$

It will therefore be an $m \times k$ matrix, where $k \leq m$. The meaning of this in terms of generators and relations for our module is that V is generated by m elements

$v_1, \ldots, v_m$, and that

$$(6.6) \qquad d_1v_1 = 0,\ d_2v_2 = 0, \ldots, d_kv_k = 0$$

forms a complete set of relations among these generators.

For $j = 1, \ldots, k$, let us denote by C_j the cyclic subgroup generated by v_j. Let L be the subgroup generated by the remaining generators $v_{k+1}, \ldots, v_m$. Since the columns of (6.5) are a complete set of relations, there is no relation involving these last $m - k$ generators. Therefore L is a free abelian group of rank $m - k$. We now verify that $V = C_1 \oplus \cdots \oplus C_k \oplus L$ and that C_j is a cyclic group of order d_j. First, since V is generated by the v_i and since each of the v_i is included in one of the summands, it is clear that V is the sum of these subgroups. Next, suppose that we have a relation, say

$$z_1 + \cdots + z_k + w = 0,$$

where $z_j \in C_j$ and $w \in L$. Since C_j is the cyclic group generated by v_j, we can write $z_j = r_jv_j$ for some integer r_j. Similarly, we may write $w = r_{k+1}v_{k+1} + \cdots + r_mv_m$ for some integers r_j. Then the relation has the form

$$r_1v_1 + \cdots + r_mv_m = 0.$$

Since the columns of (6.5) form a complete set of relations, the vector $(r_1, \ldots, r_m)^t$ is a linear combination of these columns. So $r_j = 0$ if $j > k$, which implies that $w = 0$. In addition, r_j must be divisible by d_j if $j \leq k$, say $r_j = d_js_j$. Then $z_j = s_jd_jv_j = 0$. Thus the relation was trivial, and this shows that the subgroups are independent. It also shows that the order of the cyclic group C_j is d_j. So we have $V = C_{d_1} \oplus \cdots \oplus C_{d_k} \oplus L$, as required. □

A finite abelian group is finitely generated, so as stated above the Structure Theorem decomposes a finite abelian group into a direct sum of finite cyclic groups, in which the order of each summand divides the next. The free abelian summand is zero in this case. It is sometimes convenient to decompose the cyclic groups further, into cyclic groups of prime power order. This decomposition is based on Proposition (8.4) of Chapter 2, which we restate here:

(6.7) Let r, s be relatively prime integers. The cyclic group C_{mn} of order rs is the direct sum of cyclic subgroups of orders r and s. □

Combining this lemma with the Structure Theorem yields the following:

(6.8) **Corollary.** *Structure Theorem, alternate form:* Every finitely generated abelian group is a direct sum of cyclic groups of prime power orders and of a free abelian group. □

It is natural to ask whether the orders of the cyclic subgroups which decompose a given finite abelian group are uniquely determined by the group. If the order of V

is a product of distinct primes, there is no problem. For example, if the order is 30, then V must be isomorphic to $C_2 \oplus C_3 \oplus C_5$. But can the same group be both $C_2 \oplus C_2 \oplus C_4$ and $C_4 \oplus C_4$? It is not difficult to show that this is impossible by counting elements of orders 1 or 2. The group $C_4 \oplus C_4$ contains four such elements, while $C_2 \oplus C_2 \oplus C_4$ contains eight. This counting method will always work.

(6.9) **Theorem.** *Uniqueness for the Structure Theorem:*

(a) Suppose that a finite abelian group V is a direct sum of cyclic groups $C_{d_1} \oplus \cdots \oplus C_{d_k}$ where $d_1 | d_2 | \dots$. The integers d_j are determined by the group V.

(b) The same is true if the decomposition is into prime power orders, that is, if each d_j is the power of a prime.

We leave the proof as an exercise. □

The counting of elements is simplified notationally by representing a direct sum as a product. Let R be a ring. The *direct product* of R-modules $W_1, \dots, W_k$ is the product set $W_1 \times \cdots \times W_k$ of k-tuples:

$$W_1 \times \cdots \times W_k = \{(w_1, \dots, w_k) \mid w_i \in W_i\}. \tag{6.10}$$

It is made into a module by vector addition and scalar multiplication:

$$(w_1, \dots, w_k) + (w_1', \dots, w_k') = (w_1 + w_1', \dots, w_k + w_k'), \quad r(w_1, \dots, w_k) = (rw_1, \dots, rw_k).$$

Verification of the axioms for a module is routine.

Direct products and direct sums are isomorphic, as the following proposition shows:

(6.11) **Proposition.** Let $W_1, \dots, W_k$ be submodules of an R-module V.

(a) The map $\sigma: W_1 \times \cdots \times W_k \longrightarrow V$ defined by

$$\sigma(w_1, \dots, w_k) = w_1 + \cdots + w_k$$

is a homomorphism of R-modules, and its image is the sum $W_1 + \cdots + W_k$.

(b) The homomorphism σ is an isomorphism if and only if V is the direct sum of the submodules W_i.

We have seen similar arguments several times before, so we omit the proof. Note that the second part of the proposition is analogous to the statement that the map (2.5) $R^k \longrightarrow V$ defined by a set $(v_1, \dots v_k)$ is bijective if and only if this set is a basis. □

Since a cyclic group C_d of order d is isomorphic to the standard cyclic group $\mathbb{Z}/(d)$, we can use Proposition (6.11) to restate the Structure Theorem as follows:

(6.12) **Theorem.** *Product version of the Structure Theorem:* Every finitely generated abelian group V is isomorphic to a direct product of cyclic groups

$$\mathbb{Z}/(d_1) \times \cdots \times \mathbb{Z}/(d_k) \times \mathbb{Z}^r,$$

where d_i, r are integers. There is a decomposition in which each d_i divides the next and one in which each d_i is a prime power. □

This classification of abelian groups carries over to Euclidean domains without essential change. Since a Euclidean domain R is noetherian, any finitely generated R-module V has a presentation matrix (5.6), and by the diagonalization theorem (4.6) there is a presentation matrix A which is diagonal.

To carry along the analogy with abelian groups, we define a *cyclic R-module V* to be one which is generated by a single element v. This is equivalent with saying that V is isomorphic to a quotient module R/I, where I is the ideal of R elements α such that $\alpha v = 0$. Namely, the map $\varphi: R \longrightarrow V$ sending $r \rightsquigarrow rv$ is a surjective homomorphism of modules because v generates V, and the kernel of φ, the module of relations, is a submodule of R, an ideal I (1.3). So V is isomorphic to R/I by the First Isomorphism Theorem. Conversely, if $R/I \longrightarrow V$ is an isomorphism, the image of 1 will generate V. If R is a Euclidean domain, then the ideal I will be principal, so V will be isomorphic to $R/(\alpha)$ for some $\alpha \in R$. In this case the module of relations will also be generated by a single element.

Proceeding as in the case of abelian groups, one proves the following theorem:

(6.13) **Theorem.** *Structure Theorem for modules over Euclidean domains:*

(a) Let V be a finitely generated module over a Euclidean domain R. Then V is a direct sum of cyclic modules C_j and a free module L. Equivalently, there is an isomorphism

$$\varphi: V \longrightarrow R/(d_1) \times \cdots \times R/(d_k) \times R^r$$

of V with a direct product of cyclic modules $R/(d_i)$ and a free module R^r, where r is nonnegative, the elements $d_1, \ldots d_k$ are not units and not zero, and d_i divides d_{i+1} for each $i = 1, \ldots, k-1$.

(b) The same assertion as (*a*), except that the condition that d_i divides d_{i+1} is replaced by this: Each d_i is a power of a prime element of R. Thus V is isomorphic to a product of the form

$$R/(p_1^{e_1}) \times \cdots \times R/(p_n^{e_n}) \times R^r,$$

with repetitions of primes allowed. □

For example, consider the $F[t]$-*module* V presented by the matrix A of Example (4.7). According to (5.12), it is also presented by the diagonal matrix

$$A' = \begin{bmatrix} 1 & \\ & (t-1)^2(t-2) \end{bmatrix},$$

and we can drop the first row and column from this matrix (5.12). So V is presented by the 1×1 matrix $[g]$, where $g(t) = (t-1)^2(t-2)$. This means that V is a cyclic module, isomorphic to $F[t]/(g)$. Since g has two relatively prime factors, V can be further decomposed. It is isomorphic to the direct product of two cyclic modules

(6.14) $$V \approx F[t]/(g) \approx [F[t]/(t-1)^2] \times [F[t]/(t-2)]. \square$$

With slightly more work, Theorem (6.13) can be extended to modules over any principal ideal domain. It is also true that the prime powers occurring in (b) are unique up to unit factors. A substitute for the counting argument which proves Theorem (6.9) must be found to prove this fact. We will not carry out the proof.

7. APPLICATION TO LINEAR OPERATORS

In this section we apply the theory developed in the last section in a novel way to linear operators on vector spaces over a field. This application provides a good example of the way "proof analysis" can lead to new results in mathematics. The method developed first for abelian groups is extended formally to modules over Euclidean domains. Then it is applied to a concrete new situation in which the ring is a polynomial ring. This was not the historical development. The theories for abelian groups and for linear operators were developed independently and were tied together later. But it is striking that the two cases, abelian groups and linear operators, can be formally analogous and yet end up looking so different when the same theory is applied to them.

The key observation which allows us to proceed is that if we are given a linear operator

(7.1) $$T: V \longrightarrow V$$

on a vector space over a field F, then we can use this operator to make V into a module over the polynomial ring $F[t]$. To do so, we have to define multiplication of a vector v by a polynomial $f(t) = a_n t^n + \cdots + a_1 t + a_0$. We set

(7.2) $$f(t)v = a_n T^n(v) + a_{n-1}T^{n-1}(v) + \cdots + a_1 T(v) + a_0 v.$$

The right side can be written as $[f(T)](v)$, where $f(T)$ denotes the linear operator $a_n T^n + a_{n-1}T^{n-1} + \cdots + a_1 T + a_0 I$ obtained by substituting T for t. The brackets have been added only for clarity. With this notation, we obtain the formulas

(7.3) $$tv = T(v) \quad \text{and} \quad f(t)v = [f(T)](v).$$

The fact that rule (7.2) makes V into an $F[t]$-module is easy to verify. The formulas (7.3) may appear tautological. They raise the question of why we need a new symbol t. But remember that $f(t)$ is a formal polynomial, while $f(T)$ denotes a certain linear operator.

Conversely, let V be an $F[t]$-module. Then scalar multiplication of elements of V by a polynomial $f(t)$ is defined. In particular, we are given a rule for multiplying

by the constant polynomials, the elements of F. If we keep the rule for multiplying by constants but forget for the moment about multiplication by nonconstant polynomials, then the axioms (1.1) show that V becomes a vector space over F. Next, we can multiply elements of V by the polynomial t. Let us denote the operation of multiplication by t on V as T. Thus T is the map

$$(7.4) \qquad T: V \longrightarrow V, \quad \text{defined by } T(v) = tv.$$

This map is a *linear operator* on V, when it is considered as a vector space over F. For $t(v + v') = tv + tv'$ by the distributive law (1.1), and hence $T(v + v') = T(v) + T(v')$. And if $c \in F$, then $tcv = ctv$ by the associative law (1.1) and the commutative law in $F[t]$; hence $T(cv) = cT(v)$. So an $F[t]$-module V provides us with a linear operator on a vector space.

The operations we have described, going from linear operators to modules and back, are inverses of each other:

(7.5) *Linear operator on an F-vector space and $F[t]$-module are equivalent concepts.*

We will want to apply this observation to finite-dimensional vector spaces, but let us note in passing the linear operator which corresponds to the free $F[t]$-module $F[t]$ of rank 1. We know that $F[t]$ is infinite-dimensional when it is considered as a vector space over F. The monomials $(1, t, t^2, \ldots)$ form a basis, and we can use this basis to identify $F[t]$ with the space Z of infinite F-vectors, as in Chapter 10 (2.8):

$$Z = \{(a_0, a_1, a_2, \ldots) \mid a_i \in F \text{ and only finitely many } a_i \text{ are nonzero}\}.$$

Multiplication by t on $F[t]$ corresponds to the *shift operator* T:

$$(a_0, a_1, a_2, \ldots) \rightsquigarrow (0, a_0, a_1, a_2, \ldots).$$

Thus, up to isomorphism, the free $F[t]$-module of rank 1 corresponds to the shift operator on the space Z.

We now begin our application to linear operators. Given a linear operator T on a vector space V over F, we may also view V as an $F[t]$-module. Let us suppose that V is finite-dimensional as a vector space, say of dimension n. Then it is certainly finitely generated as a module, and hence it has a presentation matrix. There is some danger of confusion here because there are two matrices around: the presentation matrix for the module V, and the matrix of the linear operator T. The presentation matrix is an $r \times s$ matrix with polynomial entries, where r is the number of chosen generators for the module and s is the number of relations. On the other hand, the matrix of the linear operator is an $n \times n$ matrix whose entries are scalars, where n is the dimension of V as a vector space. Both matrices contain the information needed to describe the module and the linear operator.

Regarding V as an $F[t]$-module, we can apply Theorem (6.13) to conclude that V is a direct sum of cyclic submodules, say

$$V = W_1 \oplus \cdots \oplus W_k,$$

where W_i is isomorphic to $F[t]/(p_i^{e_i})$, $p_i(t)$ being an irreducible polynomial in $F[t]$. There is no free summand, because we are assuming that V is finite-dimensional.

We have two tasks: to interpret the meaning of the direct sum decomposition for the linear operator T, and to describe the linear operator when the module is cyclic. It will not be surprising that the direct sum decomposition gives us a block decomposition of the matrix of T, when a suitable basis is chosen. The reason is that each of the subspaces W_i is T-invariant, because W_i is an $F[t]$-submodule. Multiplication by t carries W_i to itself, and t operates on V as the linear operator T. We choose bases $\mathbf{B}_i$ for the subspaces W_i. Then the matrix of T with respect to the basis $\mathbf{B} = (\mathbf{B}_1,\ldots,\mathbf{B}_k)$ has the desired block form [Chapter 4 (3.8)].

Next, let W be a cyclic $F[t]$-module. Then W is generated as a *module* by a single element w; in other words, every element of W can be written in the form

$$g(t)w = b_r t^r w + \cdots + b_1 t w + b_0 w,$$

where $g(t) = b_r t^r + \cdots + b_1 t + b_0 \in F[t]$. This implies that the elements $w, tw, t^2w, \ldots$ span W as a vector space. In terms of the linear operator, W is spanned by the vectors $w, T(w), T^2(w), \ldots$.

Various relations between properties of an $F[t]$-module and the corresponding linear operator are summed up in the table below.

(7.6) **Dictionary.**

$F[t]$-module	*Linear operator T*
multiplication by t	operation of T
free module of rank 1	shift operator
cyclic module generated by v	vector space spanned by $v, T(v), T^2(v), \ldots$
submodule	T-invariant subspace
direct sum of submodules	direct sum of T-invariant subspaces

Let us now compute the matrix of a linear operator T on a vector space which corresponds to a cyclic $F[t]$-module. Since every ideal of $F[t]$ is principal, such a module will be isomorphic to a module of the form

$$W = F[t]/(f), \tag{7.7}$$

where $f = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$ is a polynomial in $F[t]$. Let us use the symbol w_0 to denote the residue of 1 in W. This is our chosen generator for the module. Then the relation $fw_0 = 0$ holds, and f generates the module of relations.

The elements $w_0, tw_0, \ldots, t^{n-1}w_0$ form a basis for $F[t]/(f)$ [see Chapter 10 (5.7)]. Let us denote this basis by $w_i = t^i w_0$. Then

$$tw_0 = w_1, \quad tw_1 = w_2, \ldots, \quad tw_{n-2} = w_{n-1},$$

and also $fw_0 = 0$. This last relation can be rewritten using the others in order to determine the action of t on w_{n-1}:

$$(t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0)w_0 = tw_{n-1} + a_{n-1}w_{n-1} + \cdots + a_1 w_1 + a_0 w_0 = 0.$$

Since T acts as multiplication by t, we have

$$T(w_0) = w_1, \quad T(w_1) = w_2, \ldots, \quad T(w_{n-2}) = w_{n-1},$$

and

$$T(w_{n-1}) = -a_{n-1}w_{n-1} - \cdots - a_1 w_1 - a_0 w_0.$$

This determines the matrix of T. It has the form illustrated below for various values of n:

$$(7.8) \qquad [-a_0], \begin{bmatrix} 0 & -a_0 \\ 1 & -a_1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & -a_0 \\ 1 & 0 & -a_1 \\ 0 & 1 & -a_2 \end{bmatrix}, \ldots, \begin{bmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & 1 & 0 & & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & 0 & \vdots \\ & & & 1 & -a_{n-1} \end{bmatrix}.$$

(7.9) **Theorem.** Let T be a linear operator on a finite-dimensional vector space V over a field F. There is a basis for V with respect to which the matrix of T is made up of blocks of the type (7.8). □

Such a form for the matrix of a linear operator is called a *rational canonical form*. It isn't particularly nice, but it is the best form available for an arbitrary field.

For example, the module (6.14) is a direct sum of two modules. Its rational canonical form is

$$(7.10) \qquad \left[\begin{array}{cc|c} & -1 & \\ 1 & 2 & \\ \hline & & 2 \end{array}\right].$$

We now consider more carefully the case that F is the field of complex numbers. Every irreducible polynomial in $\mathbb{C}[t]$ is linear, $p(t) = t - \alpha$, so according to Theorem (6.12), every finite-dimensional $\mathbb{C}[t]$-module is a direct sum of submodules isomorphic to ones of the form

$$(7.11) \qquad W = \mathbb{C}[t]/(t - \alpha)^n.$$

We let w_0 denote the residue of 1 in W as before, but we make a different choice of basis for W this time, setting $w_i = (t-\alpha)^i w_0$. Then

$$(t-\alpha)w_0 = w_1, \quad (t-\alpha)w_1 = w_2, \ldots, \quad (t-\alpha)w_{n-2} = w_{n-1}, \quad \text{and } (t-\alpha)w_{n-1} = 0.$$

We replace t by T and solve, obtaining

$$Tw_i = w_{i+1} + \alpha w_i,$$

for $i = 0, \ldots, n-2$, and

$$Tw_{n-1} = \alpha w_{n-1}.$$

The matrix of T has the form

$$(7.12)\qquad [\alpha], \begin{bmatrix} \alpha & 0 \\ 1 & \alpha \end{bmatrix}, \begin{bmatrix} \alpha & 0 & 0 \\ 1 & \alpha & 0 \\ 0 & 1 & \alpha \end{bmatrix}, \ldots, \begin{bmatrix} \alpha & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & \alpha \end{bmatrix}.$$

These matrices are called *Jordan blocks*. Thus we obtain the following theorem:

(7.13) **Theorem.** Let $T\colon V \longrightarrow V$ be a linear operator on a finite-dimensional complex vector space. There is a basis of V such that the matrix of T with respect to this basis is made up of Jordan blocks. □

Such a matrix is said to be in *Jordan form*, or to be a *Jordan matrix*. Note that it is lower triangular, so the diagonal entries are its eigenvalues. Jordan form is much nicer than rational canonical form.

It is not hard to show that every Jordan block has a unique eigenvector.

Given any square complex matrix A, the theorem asserts that PAP^{-1} is in Jordan form for some invertible matrix P. We often refer to PAP^{-1} as "the Jordan form for A." It is unique up to permutation of the blocks, because the terms in the direct sum decomposition are unique, though we have not proved this.

The Jordan form of the module (6.14) is made up of two Jordan blocks:

$$(7.14)\qquad \left[\begin{array}{cc|c} 1 & & \\ 1 & 1 & \\ \hline & & 2 \end{array}\right].$$

One important application of Jordan form is to the explicit solution of systems of a first-order linear differential equation

$$(7.15)\qquad \frac{dX}{dt} = AX.$$

As we saw in Chapter 4 (7.11), the problem of solving this equation reduces easily to solving the equation $\dfrac{dX}{dt} = \tilde{A}X$, where $\tilde{A} = PAP^{-1}$ is any similar matrix. So provided that we can determine the Jordan form $\tilde{A}$ of the given matrix A, it is enough to solve the resulting system. This in turn reduces to the case of a single Jordan block. One example of a 2×2 Jordan block was computed in Chapter 4 (8.18).

The solutions for an arbitrary $k \times k$ Jordan block A can be determined by computing the matrix exponential. We denote by N the $k \times k$ matrix obtained by substituting $\alpha = 0$ into (7.12). Then $N^k = 0$. Hence

$$e^{Nt} = I + Nt/1! + \cdots + N^{k-1}t^{k-1}/(k-1)!.$$

This is a lower triangular matrix which is constant on diagonal bands and whose entries on the ith diagonal band below the diagonal are $t^i/i!$. Since N and αI

commute,

$$e^{At} = e^{\alpha t}e^{Nt} = e^{\alpha t}(I + Nt/1! + \cdots + N^{k-1}t^{k-1}/(k-1)!).$$

Thus if A is the matrix

$$A = \begin{bmatrix} 3 & & \\ 1 & 3 & \\ & 1 & 3 \end{bmatrix},$$

then

$$e^{At} = \begin{bmatrix} e^{3t} & & \\ & e^{3t} & \\ & & e^{3t} \end{bmatrix} \begin{bmatrix} 1 & & \\ t & 1 & \\ \frac{1}{2}t^2 & t & 1 \end{bmatrix} = \begin{bmatrix} e^{3t} & & \\ te^{3t} & e^{3t} & \\ \frac{1}{2}t^2e^{3t} & te^{3t} & e^{3t} \end{bmatrix}.$$

Theorem (8.14) of Chapter 4 tells us that the columns of this matrix form a basis for the space of solutions of the differential equation (7.15).

Computing the Jordan form of a given matrix requires finding the roots of its characteristic polynomial $p(t)$. If the roots $\alpha_1, \ldots, \alpha_n$ are distinct, the Jordan form is diagonal:

$$\begin{bmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_k \end{bmatrix}.$$

Suppose that the root $\alpha_1 = \alpha$ is an r-fold root of $p(t)$. Then there are various possibilities for the part of the Jordan matrix with diagonal entries α. Here are the possibilities for small r:

$$r = 1: [\alpha]; \quad r = 2: \begin{bmatrix} \alpha & \\ 1 & \alpha \end{bmatrix}, \left[\begin{array}{c|c} \alpha & \\ \hline & \alpha \end{array}\right];$$

$$r = 3: \begin{bmatrix} \alpha & & \\ 1 & \alpha & \\ & 1 & \alpha \end{bmatrix}, \left[\begin{array}{cc|c} \alpha & & \\ 1 & \alpha & \\ \hline & & \alpha \end{array}\right], \left[\begin{array}{c|c|c} \alpha & & \\ \hline & \alpha & \\ \hline & & \alpha \end{array}\right];$$

$$r = 4: \begin{bmatrix} \alpha & & & \\ 1 & \alpha & & \\ & 1 & \alpha & \\ & & 1 & \alpha \end{bmatrix}, \left[\begin{array}{ccc|c} \alpha & & & \\ 1 & \alpha & & \\ & 1 & \alpha & \\ \hline & & & \alpha \end{array}\right], \left[\begin{array}{cc|cc} \alpha & & & \\ 1 & \alpha & & \\ \hline & & \alpha & \\ & & 1 & \alpha \end{array}\right], \left[\begin{array}{cc|c|c} \alpha & & & \\ 1 & \alpha & & \\ \hline & & \alpha & \\ \hline & & & \alpha \end{array}\right],$$

$$\left[\begin{array}{c|c|c|c} \alpha & & & \\ \hline & \alpha & & \\ \hline & & \alpha & \\ \hline & & & \alpha \end{array}\right].$$

They can be distinguished by computing eigenvectors of certain operators related to T. The space of solutions to the system of equations

$$(A - \alpha I)x = 0$$

is the space of eigenvectors of A with eigenvalue α. One can solve this system explicitly, given A and α. If $r = 4$, the dimensions of the solution space in the five cases shown above are $1, 2, 2, 3, 4$ respectively, because one eigenvector is associated to each block. So this dimension distinguishes all cases except the second and third. These remaining two cases can be distinguished by the matrix $(A - \alpha I)^2$. It is zero in case three and not zero in case two.

It can be shown that the dimensions of the null spaces of the operators $(A - \alpha I)^\nu$, $\nu = 1, 2, \ldots, r/2$, distinguish the Jordan forms in all cases.

8. FREE MODULES OVER POLYNOMIAL RINGS

The structures of modules over a ring become increasingly complicated with increasing complication of the ring. It is even difficult to determine whether or not an explicitly presented module is free. In this section we describe, without proof, a theorem which characterizes free modules over polynomial rings. This theorem was proved by Quillen and Suslin in 1976.

Let $R = \mathbb{C}[x_1, \ldots, x_k]$ be the polynomial ring in k variables, and let V be a finitely generated R-module. We choose a presentation matrix A for the module. The entries of A will be polynomials $a_{ij}(x)$, and if A is an $m \times n$ matrix, then V is isomorphic to the cokernel R^m/AR^n of multiplication by A on R-vectors. We can evaluate the matrix entries $a_{ij}(x)$ at any point $p = (p_1, \ldots p_k)$ of $\mathbb{C}^k$, obtaining a complex matrix $A(p)$ whose i,j-entry is $a_{ij}(p)$.

(8.1) **Theorem.** Let V be a finitely generated module over the polynomial ring $\mathbb{C}[x_1, \ldots, x_k]$, and let A be an $m \times n$ presentation matrix for V. Denote by $A(p)$ the evaluation of A at a point $p \in \mathbb{C}^k$. Then V is a free module of rank r if and only if $A(p)$ has rank $m - r$ for every point p. □

The proof of this theorem requires background which we don't have. However, we can easily see how to use it to determine whether or not a given module is free. For example, consider the polynomial ring in two variables: $R = \mathbb{C}[x, y]$. Let V be the module presented by the 4×2 matrix

$$A = \begin{bmatrix} 1 & x \\ y & x+3 \\ x & y \\ x^2 & y^2 \end{bmatrix}. \qquad (8.2)$$

So V has four generators and two relations. Let p be a point $(a, b) \in \mathbb{C}^2$. The two

columns of the matrix A_p are

$$v_1 = (1, b, a, a^2)^t, \quad v_2 = (a, a+3, b, b^2)^t.$$

It is not hard to show that these two vectors are linearly independent for every choice of a, b, from which it follows that the rank of $A(p)$ is 2 for every point (a, b). For suppose that the vectors are dependent: $v_2 = cv_1$, or vice versa. Then the first coordinates show that $v_2 = av_1$, hence

$$(8.3) \qquad a+3 = ab, \quad b = a^2, \quad b^2 = a^3.$$

These equations have no common solutions. By Theorem (8.1), V is a free module of rank 2.

We can get an intuitive understanding for this theorem by considering the vector space $V_p = \mathbb{C}^m/A(p)\,\mathbb{C}^n$ which is presented by the complex matrix $A(p)$. It is natural to think of this vector space as a kind of "evaluation of the module V at the point p," and it can be shown that V_p is essentially independent of the choice of the presentation matrix. Therefore we can use the module V to associate a vector space V_p to every point $p \in \mathbb{C}^k$. If we imagine moving the point p about, then the vector space V_p will vary in a continuous way, providing that its dimension does not jump around. This is because the matrix $A(p)$ presenting V_p depends continuously on p. Families of vector spaces of constant dimension, parametrized by a topological space, are called *vector bundles*. The module is free if and only if the family of vector spaces V_p forms a vector bundle.

"Par une déformation coutumière aux mathématiciens, je me'en tenais au point de vue trop restreint.

Jean-Louis Verdier

EXERCISES

1. The Definition of a Module

1. Let R be a ring, considered as an R-module. Determine all module homomorphisms $\varphi: R \longrightarrow R$.
2. Let W be a submodule of an R-module V. Prove that the additive inverse of an element of W is in W.
3. Let $\varphi: V \longrightarrow W$ be a homomorphism of modules over a ring R, and let V', W' be submodules of V, W respectively. Prove that $\varphi(V')$ is a submodule of W and that $\varphi^{-1}(W')$ is a submodule of V.
4. (a) Let V be an abelian group. Prove that if V has a structure of $\mathbb{Q}$-module with its given law of composition as addition, then this structure is uniquely determined.
   (b) Prove that no finite abelian group has a $\mathbb{Q}$-module structure.
5. Let $R = \mathbb{Z}[\alpha]$, where α is an algebraic integer. Prove that for any integer m, R/mR is finite, and determine its order.

6. A module is called *simple* if it is not the zero module and if it has no proper submodule.
   (a) Prove that any simple module is isomorphic to R/M, where M is a maximal ideal.
   (b) Prove *Schur's Lemma:* Let $\varphi: S \longrightarrow S'$ be a homomorphism of simple modules. Then either φ is zero, or else it is an isomorphism.
7. The *annihilator* of an R-module V is the set $I = \{r \in R \mid rV = 0\}$.
   (a) Prove that I is an ideal of R.
   (b) What is the annihilator of the $\mathbb{Z}$-module $\mathbb{Z}/(2) \times \mathbb{Z}/(3) \times \mathbb{Z}/(4)$? of the $\mathbb{Z}$-module $\mathbb{Z}$?
8. Let R be a ring and V an R-module. Let E be the set of *endomorphisms* of V, meaning the set of homomorphisms from V to itself. Prove that E is a noncommutative ring, with composition of functions as multiplication and with addition defined by $[\varphi + \psi](m) = \varphi(m) + \psi(m)$.
9. Prove that the ring of endomorphisms of a simple module is a field.
10. Determine the ring of endomorphisms of the R-module (a) R and (b) R/I, where I is an ideal.
11. Let $W \subset V \subset U$ be R-modules.
    (a) Describe natural homomorphisms which relate the three quotient modules U/W, U/V, and V/W.
    (b) Prove the *Third Isomorphism Theorem:* $U/V \approx (U/W)/(V/W)$.
12. Let V, W be submodules of a module U.
    (a) Prove that $V \cap W$ and $V + W$ are submodules.
    (b) Prove the *Second Isomorphism Theorem:* $(V + W)/W$ is isomorphic to $V/(V \cap W)$.
13. Let V be an R-module, defined as in (1.1). If the ring R is not commutative, it is not a good idea to define $vr = rv$. Explain.

2. Matrices, Free Modules, and Bases

1. Let $R = \mathbb{C}[x, y]$, and let M be the ideal of R generated by the two elements (x, y). Prove or disprove: M is a free R-module.
2. Let A be an $n \times n$ matrix with coefficients in a ring R, let $\varphi: R^n \longrightarrow R^n$ be left multiplication by A, and let $d = \det A$. Prove or disprove: The image of φ is equal to dR^n.
3. Let I be an ideal of a ring R. Prove or disprove: If R/I is a free R-module, then $I = 0$.
4. Let R be a ring, and let V be a free R-module of finite rank. Prove or disprove:
   (a) Every set of generators contains a basis.
   (b) Every linearly independent set can be extended to a basis.
5. Let I be an ideal of a ring R. Prove that I is a free R-module if and only if it is a principal ideal, generated by an element α which is not a zero divisor in R.
6. Prove that a ring R such that every finitely generated R-module is free is either a field or the zero ring.
7. Let A be the matrix of a homomorphism $\varphi: \mathbb{Z}^n \longrightarrow \mathbb{Z}^m$ between free modules.
   (a) Prove that φ is injective if and only if the rank of A is n.
   (b) Prove that φ is surjective if and only if the greatest common divisor of the determinants of the $m \times m$ minors of A is 1.
8. Reconcile the definition of free abelian group given in Section 2 with that given in Chapter 6, Section 8.

3. The Principle of Permanence of Identities

1. In each case, decide whether or not the principle of permanence of identities allows the result to be carried over from the complex numbers to an arbitrary commutative ring.
   (a) the associative law for matrix multiplication
   (b) Cayley–Hamilton Theorem
   (c) Cramer's Rule
   (d) product rule, quotient rule, and chain rule for differentiation of polynomials
   (e) the fact that a polynomial of degree n has at most n roots
   (f) Taylor's expansion of a polynomial
2. Does the principle of permanence of identities show that $\det AB = \det A \det B$ when the entries of the matrices are in a noncommutative ring R?
3. In some cases, it may be convenient to verify an identity only for the real numbers. Does this suffice?
4. Let R be a ring, and let A be a 3×3 R-matrix in $SO_3(R)$, that is, such that $A^t A = I$ and $\det A = 1$. Does the principle of permanence of identities show that A has an eigenvector in R^3 with eigenvalue 1?

4. Diagonalization of Integer Matrices

1. Reduce each matrix below to diagonal form by integer row and column operations.
   (a) $\begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ (c) $\begin{bmatrix} 3 & 1 & -4 \\ 2 & -3 & 1 \\ -4 & 6 & -2 \end{bmatrix}$
   (d) In the first case, let $V = \mathbb{Z}^2$ and let $L = AV$. Draw the sublattice L, and find commensurable bases of V and L.
2. Let A be a matrix whose entries are in the polynomial ring $F[t]$, and let A' be obtained from A by polynomial row and column operations. Relate $\det A$ to $\det A'$.
3. Determine integer matrices P^{-1}, Q which diagonalize the matrix $A = \begin{bmatrix} 4 & 7 & 2 \\ 2 & 4 & 6 \end{bmatrix}$.
4. Let $d_1, d_2, \ldots$ be the integers referred to in Theorem (4.3).
   (a) Prove that d_1 is the greatest common divisor of the entries a_{ij} of A.
   (b) Prove that $d_1 d_2$ is the greatest common divisor of the determinants of the 2×2 minors of A.
   (c) State and prove an extension of (a) and (b) to d_i for arbitrary i.
5. Determine all integer solutions to the system of equations $AX = 0$, when $A = \begin{bmatrix} 4 & 7 & 2 \\ 2 & 4 & 6 \end{bmatrix}$.
6. Find a basis for the following submodules of $\mathbb{Z}^3$.
   (a) The module generated by $(1, 0, -1)$, $(2, -3, 1)$, $(0, 3, 1)$, $(3, 1, 5)$.
   (b) The module of solutions of the system of equations $x + 2y + 3z = 0$, $x + 4y + 9z = 0$.
7. Prove that the two matrices $\begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}$ and $\begin{bmatrix} & -1 \\ 1 & \end{bmatrix}$ generate the group $SL_2(\mathbb{Z})$ of integer matrices with determinant 1.

8. Prove that the group $SL_n(\mathbb{Z})$ is generated by elementary integer matrices of the first type.

9. Let α, β, γ be complex numbers, and let $A = \{\ell\alpha + m\beta + n\gamma \mid \ell, m, n, \in \mathbb{Z}\}$ be the subgroup of $\mathbb{C}^+$ they generate. Under what conditions is A a lattice in $\mathbb{C}$?

10. Let $\varphi: \mathbb{Z}^k \longrightarrow \mathbb{Z}^k$ be a homomorphism given by multiplication by an integer matrix A. Show that the image of φ is of finite index if and only if A is nonsingular and that if so, then the index is equal to $|\det A|$.

11. **(a)** Let $A = (a_1, \ldots, a_n)^t$ be an integer column vector. Use row reduction to prove that there is a matrix $P \in GL_n(\mathbb{Z})$ such that $PA = (d, 0, \ldots, 0)^t$, where d is the greatest common divisor of $a_1, \ldots, a_n$.

(b) Prove that if $d = 1$, then A is the first column of a matrix of $M \in SL_n(\mathbb{Z})$.

5. Generators and Relations for Modules

1. In each case, identify the abelian group which has the given presentation matrix:

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \end{bmatrix}, [2 \quad 0 \quad 0], \begin{bmatrix} 1 & \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 4 \\ 1 & 4 \end{bmatrix}, \begin{bmatrix} 4 & 6 \\ 2 & 3 \end{bmatrix}.$$

2. Find a ring R and an ideal I of R which is not finitely generated.

3. Prove that existence of factorizations holds in a noetherian integral domain.

4. Let $V \subset \mathbb{C}^n$ be the locus of zeros of an infinite set of polynomials $f_1, f_2, f_3, \ldots$. Prove that there is a finite subset of these polynomials whose zeros define the same locus.

5. Let S be a subset of $\mathbb{C}^n$. Prove that there is a finite set of polynomials $(f_1, \ldots, f_k)$ such that any polynomial which vanishes identically on S is a linear combination of this set, with polynomial coefficients.

6. Determine a presentation matrix for the ideal $(2, 1 + \delta)$ of $\mathbb{Z}[\delta]$, where $\delta = \sqrt{-5}$.

***7.** Let S be a subring of the ring $R = \mathbb{C}[t]$ which contains $\mathbb{C}$ and is not equal to $\mathbb{C}$. Prove that R is a finitely generated S-module.

8. Let A be the presentation matrix of a module V with respect to a set of generators $(v_1, \ldots, v_m)$. Let $(w_1, \ldots, w_r)$ be another set of elements of V, and write the elements in terms of the generators, say $w_i = \Sigma p_{ij} v_j$, $p_{ij} \in R$. Let $P = (p_{ij})$. Prove that the block matrix $\left[\begin{array}{c|c} A & -P \\ \hline 0 & I \end{array}\right]$ is a presentation matrix for V with respect to the set of generators $(v_1, \ldots, v_m; w_1, \ldots, w_r)$.

***9.** With the notation of the previous problem, suppose that $(w_1, \ldots, w_r)$ is also a set of generators of V and that B is a presentation matrix for V with respect to this set of generators. Say that $v_i = \Sigma q_{ij} w_j$ is an expression of the generators v_i in terms of the w_j.

(a) Prove that the block matrix $M = \left[\begin{array}{c|c|c|c} A & -P & I & 0 \\ \hline 0 & I & -Q & B \end{array}\right]$ presents V with respect to the generators $(v_1, \ldots, v_m; w_1, \ldots, w_r)$.

(b) Show that M can be reduced to A and to B by a sequence of operations of the form (5.12).

10. Using 9, show that any presentation matrix of a module can be transformed to any other by a sequence of operations (5.12) and their inverses.

6. The Structure Theorem for Abelian Groups

1. Find a direct sum of cyclic groups which is isomorphic to the abelian group presented by the matrix $\begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix}$.
2. Write the group generated by x, y, with the relation $3x + 4y = 0$ as a direct sum of cyclic groups.
3. Find an isomorphic direct product of cyclic groups, when V is the abelian group generated by x, y, z, with the given relations.
   (a) $3x + 2y + 8z = 0,\ 2x + 4z = 0$
   (b) $x + y = 0,\ 2x = 0,\ 4x + 2z = 0,\ 4x + 2y + 2z = 0$
   (c) $2x + y = 0,\ x - y + 3z = 0$
   (d) $2x - 4y = 0,\quad 2x + 2y + z = 0$
   (e) $7x + 5y + 2z = 0,\ 3x + 3y = 0,\ 13x + 11y + 2z = 0$
4. Determine the number of isomorphism classes of abelian groups of order 400.
5. Classify finitely generated modules over each ring.
   (a) $\mathbb{Z}/(4)$ (b) $\mathbb{Z}/(6)$ (c) $\mathbb{Z}/n\mathbb{Z}$.
6. Let R be a ring, and let V be an R-module, presented by a diagonal $m \times n$ matrix A: $V \approx R^m/AR^n$. Let $(v_1, \dots, v_m)$ be the corresponding generators of V, and let d_i be the diagonal entries of A. Prove that V is isomorphic to a direct product of the modules $R/(d_i)$.
7. Let V be the $\mathbb{Z}[i]$-module generated by elements v_1, v_2 with relations $(1 + i)v_1 + (2 - i)v_2 = 0,\quad 3v_1 + 5iv_2 = 0$. Write this module as a direct sum of cyclic modules.
8. Let $W_1, \dots, W_k$ be submodules of an R-module V such that $V = \Sigma W_i$. Assume that $W_1 \cap W_2 = 0,\ (W_1 + W_2) \cap W_3 = 0, \dots, (W_1 + W_2 + \cdots + W_{k-1}) \cap W_k = 0$. Prove that V is the direct sum of the modules $W_1, \dots, W_k$.
9. Prove the following.
   (a) The number of elements of $\mathbb{Z}/(p^e)$ whose order divides p^ν is p^ν if $\nu \le e$, and is p^e if $\nu \ge e$.
   (b) Let $W_1, \dots, W_k$ be finite abelian groups, and let u_j denote the number of elements of W_j whose order divides a given integer q. Then the number of elements of the product group $V = W_1 \times \cdots \times W_k$ whose order divides q is $u_1 \cdots u_k$.
   (c) With the above notation, assume that W_j is a cyclic group of prime power order $d_j = p^{e_j}$. Let r_1 be the number of d_j equal to a given prime p, let r_2 be the number of d_j equal to p^2, and so on. Then the number of elements of V whose order divides p^ν is p^{s_ν}, where $s_1 = r_1 + \cdots + r_k$, $s_2 = r_1 + 2r_2 + \cdots + 2r_k$, $s_3 = r_1 + 2r_2 + 3r_3 + \cdots + 3r_k$, and so on.
   (d) Theorem (6.9).

7. Application to Linear Operators

1. Let T be a linear operator whose matrix is $\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$. Is the corresponding $\mathbb{C}[t]$-module cyclic?

2. Determine the Jordan form of the matrix $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$.

3. Prove that $\begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix}$ is an idempotent matrix, and find its Jordan form.

4. Let V be a complex vector space of dimension 5, and let T be a linear operator on V which has characteristic polynomial $(t - \alpha)^5$. Suppose that the rank of the operator $T - \alpha I$ is 2. What are the possible Jordan forms for T?

5. Find all possible Jordan forms for a matrix whose characteristic polynomial is $(t + 2)^2(t - 5)^3$.

6. What is the Jordan form of a matrix whose characteristic polynomial is $(t - 2)^2(t - 5)^3$ and such that the space of eigenvectors with eigenvalue 2 is one-dimensional, while the space of eigenvectors with eigenvalue 5 is two-dimensional?

7. **(a)** Prove that a Jordan block has a one-dimensional space of eigenvectors.
(b) Prove that, conversely, if the eigenvectors of a complex matrix A are multiples of a single vector, then the Jordan form for A consists of one block.

8. Determine all invariant subspaces of a linear operator whose Jordan form consists of one block.

9. In each case, solve the differential equation $dX/dt = AX$ when A is the Jordan block given.

(a) $\begin{bmatrix} 2 & \\ 1 & 2 \end{bmatrix}$ **(b)** $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ **(c)** $\begin{bmatrix} 1 & & \\ 1 & 1 & \\ & 1 & 1 \end{bmatrix}$

10. Solve the differential equation $dX/dt = AX$ when A is **(a)** the matrix (7.14), **(b)** the matrix (7.10), **(c)** the matrix of problem 2, **(d)** the matrix of problem 3.

11. Prove or disprove: Two complex $n \times n$ matrices A, B are similar if and only if they have the same Jordan form.

12. Show that every complex $n \times n$ matrix is similar to a matrix of the form $D + N$, where D is diagonal, N is nilpotent, and $DN = ND$.

13. Let $R = F[x]$ be the polynomial ring in one variable over a field F, and let V be the R-module generated by an element v which satisfies the relation $(x^3 + 3x + 2)v = 0$. Choose a basis for V as F-vector space, and find the matrix of the operator multiplication by t with respect to this basis.

14. Let V be an $F[t]$-module, and let $\mathbf{B} = (v_1, \ldots, v_n)$ be a basis for V, as F-vector space. Let B be the matrix of T with respect to this basis. Prove that $A = tI - B$ is a presentation matrix for the module.

15. Let $p(t)$ be a polynomial over a field F. Prove that there exists an $n \times n$ matrix with entries in F whose characteristic polynomial is $p(t)$.

16. Prove or disprove: A complex matrix A such that $A^2 = A$ is diagonalizable.

17. Let A be a complex $n \times n$ matrix such that $A^k = I$ for some n. Prove that the Jordan form for A is diagonal.

18. Prove the Cayley–Hamilton Theorem, that if $p(t)$ is the characteristic polynomial of an $n \times n$ matrix A, then $p(A) = 0$.

19. The *minimal polynomial* $m(t)$ of a linear operator T on a complex vector space V is the polynomial of lowest degree such that $m(T) = 0$.

(a) Prove that the minimal polynomial divides the characteristic polynomial.

(b) Prove that every root of the characteristic polynomial $p(t)$ is also a root of the minimal polynomial $m(t)$.

(c) Prove that T is diagonalizable if and only if $m(t)$ has no multiple root.

20. Find all possible Jordan forms for 8×8 matrices whose minimal polynomial is $x^2(x-1)^3$.

21. Prove or disprove: A complex matrix A is similar to its transpose.

22. Classify linear operators on a finitely generated $F[t]$-module, dropping the assumption that the module is finite-dimensional as a vector space.

23. Prove that the ranks of $(A - \alpha I)^\nu$ distinguish all Jordan forms, and hence that the Jordan form depends only on the operator and not on the basis.

24. Show that the following concepts are equivalent:

(i) R-module, where $R = \mathbb{Z}[i]$;

(ii) abelian group V, together with a homomorphism $\varphi: V \longrightarrow V$ such that $\varphi \circ \varphi = -\text{identity}$.

25. Let $F = \mathbb{F}_p$. For which prime integers p does the additive group F^1 have a structure of $\mathbb{Z}[i]$-module? How about F^2?

26. Classify finitely generated modules over the ring $\mathbb{C}[\epsilon]$, where $\epsilon^2 = 0$.

8. Free Modules over Polynomial Rings

1. Determine whether or not the modules over $\mathbb{C}[x,y]$ presented by the following matrices are free.

(a) $\begin{bmatrix} x^2+1 & x \\ x^2y+x+y & xy+1 \end{bmatrix}$ **(b)** $\begin{bmatrix} xy-1 \\ x^2-y^2 \\ y \end{bmatrix}$ **(c)** $\begin{bmatrix} x-1 & x \\ y & y+1 \\ x & y \\ x^2 & 2y \end{bmatrix}$

2. Prove that the module presented by (8.2) is free by exhibiting a basis.

3. Following the model of the polynomial ring in one variable, describe modules over the ring $\mathbb{C}[x,y]$ in terms of real vector spaces with additional structure.

4. Let R be a ring and V an R-module. Let I be an ideal of R, and let IV be the set of finite sums $\Sigma s_i v_i$, where $s_i \in I$ and $v_i \in V$.

(a) Show how to make V/IV into an R/I-module.

(b) Let A be a presentation matrix for V, and let $\bar{A}$ denote its residue in R/I. Prove that $\bar{A}$ is a presentation matrix for V/IV.

(c) Show why the module V_p defined in the text is essentially independent of the presentation matrix.

***5.** Using exercise 9 of Section 5, prove the easy half of the theorem of Quillen and Suslin: If V is free, then the rank of $A(p)$ is constant.

6. Let $R = \mathbb{Z}[\sqrt{-5}]$, and let V be the module presented by the matrix $A = \begin{bmatrix} 2 \\ 1+\delta \end{bmatrix}$.

(a) Prove that the residue of A has rank 1 for every prime ideal P of R.

(b) Prove that V is not free.

Miscellaneous Problems

1. Let G be a lattice group, and let g be a rotation in G. Let $\overline{g}$ be the associated element of the point group $\overline{G}$. Prove that there is a basis for $\mathbb{R}^2$, not necessarily an orthonormal basis, such that the matrix of $\overline{g}$ with respect to this basis is in $GL_2(\mathbb{Z})$.

***2.** (a) Let α be a complex number, and let $\mathbb{Z}[\alpha]$ be the subring of $\mathbb{C}$ generated by α. Prove that α is an algebraic integer if and only if $\mathbb{Z}[\alpha]$ is a finitely generated abelian group.
(b) Prove that if α, β are algebraic integers, then the subring $\mathbb{Z}[\alpha, \beta]$ of $\mathbb{C}$ which they generate is a finitely generated abelian group.
(c) Prove that the algebraic integers form a subring of $\mathbb{C}$.

***3.** *Pick's Theorem:* Let Δ be the plane region bounded by a polygon whose vertices are at integer lattice points. Let I be the set of lattice points in the interior of Δ and B the set of lattice points on the boundary of Δ. If p is a lattice point, let $r(p)$ denote the fraction of 2π of the angle subtended by Δ at p. So $r(p) = 0$ if $p \notin \Delta$, $r(p) = 1$ if p is an interior point of Δ, $r(p) = \frac{1}{2}$ if p is on an edge, and so on.
(a) Prove that the area of Δ is $\sum_p r(p)$.
(b) Prove that the area is $|I| + \frac{1}{2}(|B| - 2)$ if Δ has a single connected boundary curve.

4. Prove that the integer orthogonal group $O_n(\mathbb{Z})$ is a finite group.

***5.** Consider the space $V = \mathbb{R}^k$ of column vectors as an inner product space, with the ordinary dot product $(v \cdot w) = v^t w$. Let L be a lattice in V, and define $L^* = \{w \mid (v \cdot w) \in \mathbb{Z} \text{ for all } v \in L\}$.
(a) Show that L^* is a lattice.
(b) Let $\mathbf{B} = (v_1, \dots, v_k)$ be a lattice basis for L, and let $P = [\mathbf{B}]^{-1}$ be the matrix relating this basis of V to the standard basis $\mathbf{E}$. What is the matrix A of dot product with respect to the basis $\mathbf{B}$?
(c) Show that the columns of P form a lattice basis for L^*.
(d) Show that if A is an integer matrix, then $L \subset L^*$, and $[L^* : L] = |\det A|$.

6. Let V be a real vector space having a countably infinite basis $\{v_1, v_2, v_3, \dots\}$, and let E be the ring of linear operators on V.
(a) Which infinite matrices represent linear operators on V?
(b) Describe how to compute the matrix of the composition of two linear operators in terms of the matrix of each of them.
(c) Consider the linear operators T, T' defined by the rules

$$T(v_{2n}) = v_n, \quad T(v_{2n-1}) = 0, \quad T'(v_{2n}) = 0, \quad T'(v_{2n-1}) = v_n, \quad n = 1, 2, 3, \dots.$$

Write down their matrices.
(d) We can consider $E^1 = E$ as a module over the ring E, with scalar multiplication on the left side of a vector. Show that $\{T, T'\}$ is a basis of E^1 as E-module.
(e) Prove that the free E-modules E^k, $k = 1, 2, 3 \dots$, are all isomorphic.

7. Prove that the group $\mathbb{Q}^+/\mathbb{Z}^+$ is not an infinite direct sum of cyclic groups.

8. Prove that the additive group $\mathbb{Q}^+$ of rational numbers is not a direct sum of two proper subgroups.

9. Prove that the multiplicative group $\mathbb{Q}^\times$ of rational numbers is isomorphic to the direct sum of a cyclic group of order 2 and a free abelian group with countably many generators.

10. Prove that two diagonalizable matrices are simultaneously diagonalizable, that is, that there is an invertible matrix P such that PAP^{-1} and PBP^{-1} are both diagonal, if and only if $AB = BA$.

***11.** Let A be a finite abelian group, and let $\varphi: A \longrightarrow \mathbb{C}^\times$ be a homomorphism which is not the trivial homomorphism ($\varphi(x) = 1$ for all x). Prove that $\sum_{a \in A} \varphi(a) = 0$.

12. Let A be an $m \times n$ matrix with coefficients in a ring R, and let $\varphi: R^n \longrightarrow R^m$ be left multiplication by A. Prove that the following are equivalent:
(i) φ is surjective;
(ii) the determinants of the $m \times m$ minors of A generate the unit ideal;
(iii) A has a right inverse, a matrix B with coefficients in R such that $AB = I$.

***13.** Let $(v_1, \dots, v_m)$ be generators for an R-module V, and let J be an ideal of R. Define JV to be the set of all finite sums of products av, $a \in J$, $v \in V$.
(a) Show that if $JV = V$, there is an $n \times n$ matrix A with entries in J such that $(v_1, \dots, v_m)(I - A) = 0$.
(b) With the notation of (a), show that $\det(I - A) = 1 + \alpha$, where $\alpha \in J$, and that $\det(I - A)$ annihilates V.
(c) An R-module V is called *faithful* if $rV = 0$ for $r \in R$ implies $r = 0$. Prove the *Nakayama Lemma:* Let V be a finitely generated, faithful R-module, and let J be an ideal of R. If $JV = V$, then $J = R$.
(d) Let V be a finitely generated R-module. Prove that if $MV = V$ for all maximal ideals M, then $V = 0$.

***14.** We can use a pair $x(t), y(t)$ of complex polynomials in t to define a complex path in $\mathbb{C}^2$, by sending $t \rightsquigarrow (x(t), y(t))$. They also define a homomorphism $\varphi: \mathbb{C}[x,y] \longrightarrow \mathbb{C}[t]$, by $f(x,y) \rightsquigarrow f(x(t), y(t))$. This exercise analyzes the relationship between the path and the homomorphism. Let's rule out the trivial case that $x(t), y(t)$ are both constant.
(a) Let S denote the image of φ. Prove that S is isomorphic to the quotient $\mathbb{C}[x,y]/(f)$, where $f(x,y)$ is an irreducible polynomial.
(b) Prove that t is the root of a monic polynomial with coefficients in S.
(c) Let V denote the variety of zeros of f in $\mathbb{C}^2$. Prove that for every point $(x_0, y_0) \in V$, there is a $t_0 \in \mathbb{C}$ such that $(x_0, y_0) = (x(t_0), y(t_0))$.

Chapter 13

Fields

Our difficulty is not in the proofs, but in learning what to prove.

Emil Artin

1. EXAMPLES OF FIELDS

Much of the theory of fields has to do with a pair $F \subset K$ of fields, one contained in the other. In contrast with group theory, where subgroups play an important role, we usually consider K as an extension of F; that is, F is considered to be the basic field, and K is related to it. An *extension field* of F is a field which contains F as a subfield.

Here are the three most important classes of fields.

(1.1) **Number fields.** A number field K is a subfield of $\mathbb{C}$.

Any subfield of $\mathbb{C}$ contains 1, and hence it contains the field $\mathbb{Q}$ of rational numbers. So a number field is an extension of $\mathbb{Q}$. The number fields most commonly studied are *algebraic* number fields, all of whose elements are algebraic numbers (see Chapter 10, Section 1). We studied quadratic number fields in Chapter 11.

(1.2) **Finite fields.** A field having finitely many elements is called a finite field.

If K is a finite field, then the kernel of the unique homomorphism $\varphi: \mathbb{Z} \longrightarrow K$ is a prime ideal [Chapter 11 (7.15)], and since $\mathbb{Z}$ is infinite while K is finite, the kernel is not zero. Therefore it is generated by a prime integer p. The image of φ is isomorphic to the quotient $\mathbb{Z}/(p) = \mathbb{F}_p$. So K contains a subfield isomorphic to the prime field $\mathbb{F}_p$, and therefore it can be viewed as an extension of this prime field. We will describe all finite fields in Section 6.

(1.3) **Function fields.** Certain extensions of the field $F = \mathbb{C}(x)$ of rational functions are called function fields.

Function fields play an important role in the theory of analytic funtions and in algebraic geometry. Since we haven't seen them before, we will describe them briefly here. A function field can be defined by an irreducible polynomial in two variables, say $f(x,y) \in \mathbb{C}[x,y]$. The polynomial $f(x,y) = y^2 - x^3 + x$ is a good example. Given such a polynomial f, we may study the equation

$$f(x,y) = 0 \tag{1.4}$$

analytically, using it to define y "implicitly" as a function $y(x)$ of x as we learn to do in calculus. In our example, the function defined in this way is $y = \sqrt{x^3 - x}$. This function isn't single valued; it is determined only up to sign, but that isn't a serious difficulty. We won't have an explicit expression for such a function in general, but by definition, it satisfies the equation (1.4), that is,

$$f(x,y(x)) = 0. \tag{1.5}$$

On the other hand, the equation can also be studied algebraically. Let us interpret $f(x,y)$ as a polynomial in y whose coefficients are polynomials in x. Let F denote the field $\mathbb{C}(x)$ of rational functions in x. If f is not a polynomial in x alone, then since it is irreducible in $\mathbb{C}[x,y]$, it will be an irreducible element of $F[y]$ [Chapter 11 (3.9)]. Therefore the ideal generated by f in $F[y]$ is maximal [Chapter 11 (1.6)], and $F[y]/(f) = K$ is an extension field of F.

The analysis and the algebra are related, because both the implicitly defined function $y(x)$ and the residue $\overline{y}$ of y in $F[y]/(f)$ satisfy the equation $f(x,y) = 0$. In this way, the residue of y, and indeed all elements of K, can be interpreted as functions of the variable x. Because of this, such fields are called function fields. We will discuss function fields in Section 7.

2. ALGEBRAIC AND TRANSCENDENTAL ELEMENTS

Let K be an extension of a field F, and let α be an element of K. In analogy with the definition of algebraic numbers (Chapter 10, Section 1), α is said to be *algebraic over F* if it is the root of some nonzero polynomial with coefficients in F. Since the coefficients are from a field, we may assume that the polynomial is monic, say

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0, \quad \text{with } a_i \in F. \tag{2.1}$$

An element α is called *transcendental over F* if it is not algebraic over F, that is, if it is not a root of any such polynomial.

Note that the two properties, algebraic and transcendental, depend on the given field F. For example, the complex number $2\pi i$ is algebraic over the field of real numbers but transcendental over the field of rational numbers. Also, every element α of a field K is algebraic over K, because it is the root of the polynomial $x - \alpha$, which has coefficients in K.

The two possibilities for α can be described in terms of the substitution homomorphism

$$(2.2) \qquad \varphi: F[x] \longrightarrow K, \quad \text{which maps } f(x) \rightsquigarrow f(\alpha).$$

The element α is transcendental over F if φ is injective and algebraic over F otherwise, that is, if the kernel of φ is not zero.

Assume that α is algebraic over F. Since $F[x]$ is a principal ideal domain, ker φ is generated by a single element $f(x)$, the monic polynomial of lowest degree having α as a root. Since K is a field, we know that $f(x)$ must be an irreducible polynomial [Chapter 11 (7.15)], and in fact it will be the only irreducible monic polynomial in the ideal. Every other element of the ideal is a multiple of $f(x)$. We will call this polynomial f the *irreducible polynomial for α over F.*

It is important to note that this irreducible polynomial f depends on F as well as on α, because irreducibility of a polynomial depends on the field. For example, let $F = \mathbb{Q}[i]$, and let α be the complex number $\sqrt{i} = \frac{1}{2}\sqrt{2}(1 + i)$. The irreducible polynomial for α over $\mathbb{Q}$ is $x^4 + 1$, but this polynomial factors in the field F: $x^4 + 1 = (x^2 + i)(x^2 - i)$. The irreducible polynomial for α over F is $x^2 - i$. When there are several fields around, we must be careful to make it clear to which field we refer. To say that a polynomial is irreducible is ambiguous. It is better to say that f is *irreducible over F*, or that it is an *irreducible element of $F[x]$*.

The field extension of F which is generated by an element $\alpha \in K$ will be denoted by $F(\alpha)$:

(2.3) *$F(\alpha)$ is the smallest field containing F and α.*

More generally, if $\alpha_1,\dots,\alpha_n$ are elements of an extension field K of F, then the notation $F(\alpha_1,\dots,\alpha_n)$ will stand for the smallest subfield K which contains these elements.

As in Chapter 10, we denote the *ring* generated by α over F by $F[\alpha]$. It consists of all elements of K which can be written as polynomials in α with coefficients in F:

$$(2.4) \qquad a_n\alpha^n + \cdots + a_1\alpha + a_0, \quad a_i \in F.$$

The field $F(\alpha)$ is isomorphic to the field of fractions of $F[\alpha]$. Its elements are ratios of elements of the form (2.4) [see Chapter 10 (6.7)].

(2.5) **Proposition.** If α is transcendental over F, then the map $F[x] \longrightarrow F[\alpha]$ is an isomorphism, and hence $F(\alpha)$ is isomorphic to the field $F(x)$ of rational functions. □

This simple fact has the consequence that the field extensions $F(\alpha)$ are isomorphic for all transcendental elements α, because they are all isomorphic to the field of rational functions $F(x)$. For instance, π and e are both transcendental over $\mathbb{Q}$ (though we have not proved that they are). Therefore $\mathbb{Q}(\pi)$ and $\mathbb{Q}(e)$ are isomorphic

fields, the isomorphism carrying π to e. This is rather surprising at first glance. The isomorphism is not continuous when the fields are regarded as subfields of the real numbers.

The situation is quite different if α is algebraic:

(2.6) **Proposition.**

(a) Suppose that α is algebraic over F, and let $f(x)$ be its irreducible polynomial over F. The map $F[x]/(f) \longrightarrow F[\alpha]$ is an isomorphism, and $F[\alpha]$ is a field. Thus $F[\alpha] = F(\alpha)$.

(b) More generally, let $\alpha_1, \dots, \alpha_n$ be algebraic elements of a field extension K of F. Then $F[\alpha_1, \dots, \alpha_n] = F(\alpha_1, \dots, \alpha_n)$.

Proof. Let φ be the map (2.2), with $K = F(\alpha)$. Since $f(x)$ generates ker φ, we know that $F[x]/(f)$ is isomorphic to the image of φ [Chapter 10 (3.1)], which is $F[\alpha]$. Since f is irreducible, it generates a maximal ideal [Chapter 11 (1.6)]. This shows that $F[\alpha]$ is a field. Since $F(\alpha)$ is isomorphic to the fraction field of $F[\alpha]$, it is equal to $F[\alpha]$. We leave the proof of the second part as an exercise. □

(2.7) **Proposition.** Let α be an algebraic element over F, and let $f(x)$ be its irreducible polynomial. Suppose $f(x)$ has degree n. Then $(1, \alpha, \dots, \alpha^{n-1})$ is a basis for $F[\alpha]$ as a vector space over F.

Proof. This proposition is a special case of (5.7) in Chapter 10. □

It may not be easy to tell whether or not two algebraic elements α, β generate isomorphic fields, though we can use Proposition (2.7) to give a *necessary* condition: Their irreducible polynomials over F must have the same degree, because this degree is the dimension of the field extension as an F-vector space. This is obviously not a sufficient condition. For example, all the imaginary quadratic fields studied in Chapter 11 are obtained by adjoining elements δ whose irreducible polynomials $x^2 - d$ have degree 2, but they aren't all isomorphic. On the other hand, if α is a root of $x^3 - x + 1$, then $\beta = \alpha^2$ is a root of $x^3 - 2x^2 + x - 1$. The two fields $\mathbb{Q}(\alpha)$ and $\mathbb{Q}(\beta)$ are actually equal, though if we were presented only with the two polynomials, it might take us some time to notice how they are related.

What we can describe easily are the circumstances under which there is an isomorphism

$$F(\alpha) \xrightarrow{\sim} F(\beta) \tag{2.8}$$

which fixes F and sends α to β. The following proposition is fundamental to our understanding of field extensions:

(2.9) **Proposition.** Let $\alpha \in K$ and $\beta \in L$ be algebraic elements of two extension fields of F. There is an isomorphism of fields

$$\sigma: F(\alpha) \xrightarrow{\sim} F(\beta),$$

which is the identity on the subfield F and which sends $\alpha \rightsquigarrow \beta$ if and only if the irreducible polynomials for α and β over F are equal.

Proof. Assume that $f(x)$ is the irreducible polynomial for α and for β over F. We apply Proposition (2.6), obtaining two isomorphisms

$$F[x]/(f) \xrightarrow{\varphi} F[\alpha] \quad \text{and} \quad F[x]/(f) \xrightarrow{\psi} F[\beta].$$

The composed map $\sigma = \psi\varphi^{-1}$ is the required isomorphism. Conversely, if there is an isomorphism σ sending α to β which is the identity on F, and if $f(x) \in F[x]$ is a polynomial such that $f(\alpha) = 0$, then $f(\beta) = 0$ too [see Proposition (2.11)]. Hence the two elements have the same irreducible polynomial. □

(2.10) **Definition.** Let K and K' be two extensions of the same field F. An isomorphism $\varphi: K \longrightarrow K'$ which restricts to the identity on the subfield F is called an *isomorphism of field extensions*, or an *F-isomorphism*. Two extensions K, K' of a field F are said to be *isomorphic field extensions* if there exists an F-isomorphism $\varphi: K \longrightarrow K'$.

(2.11) **Proposition.** Let $\varphi: K \longrightarrow K'$ be an isomorphism of field extensions of F, and let $f(x)$ be a polynomial with coefficients in F. Let α be a root of f in K, and let $\alpha' = \varphi(\alpha)$ be its image in K'. Then α' is also a root of f.

Proof. Say that $f(x) = a_n x^n + \cdots + a_1 x + a_0$. Then $\varphi(a_i) = a_i$ and $\varphi(\alpha) = \alpha'$. Since φ is a homomorphism, we can expand as follows:

$$\begin{aligned} 0 = \varphi(0) = \varphi(f(\alpha)) &= \varphi(a_n\alpha^n + \cdots + a_1\alpha + a_0) \\ &= \varphi(a_n)\varphi(\alpha)^n + \cdots + \varphi(a_1)\varphi(\alpha) + \varphi(a_0) \\ &= a_n\alpha'^n + \cdots + a_1\alpha' + a_0. \end{aligned}$$

This shows that α' is a root of f. □

For example, the polynomial $x^3 - 2$ is irreducible over $\mathbb{Q}$. Let α denote the real cube root of 2, and let $\zeta = e^{2\pi i/3}$ be a complex cube root of 1. The three complex roots of $x^3 - 2$ are α, $\zeta\alpha$, and $\zeta^2\alpha$. Therefore there is an isomorphism

$$\mathbb{Q}(\alpha) \xrightarrow{\sim} \mathbb{Q}(\zeta\alpha) \tag{2.12}$$

sending α to $\zeta\alpha$. In this case the elements of $\mathbb{Q}(\alpha)$ are all real numbers, but $\mathbb{Q}(\zeta\alpha)$ is not a subfield of $\mathbb{R}$. To understand the isomorphism (2.12), we must stop viewing these fields as subfields of $\mathbb{C}$ and look only at their internal algebraic structure.

3. THE DEGREE OF A FIELD EXTENSION

An extension K of a field F can always be regarded as an F-vector space. Addition is the addition law in K, and scalar multiplication of an element α of K by an element c of F is defined to be the product $c\alpha$ formed by multiplying these two elements in

K. The dimension of K as an F-vector space is called the *degree* of the field extension $F \subset K$. The degree is the simplest invariant of an extension, but though simple, it is important. It will be denoted by

(3.1) $[K : F] =$ *dimension of* K, *as an* F-*vector space.*

For example, $\mathbb{C}$ has the $\mathbb{R}$-basis $(1, i)$, so $[\mathbb{C} : \mathbb{R}] = 2$.

A field extension $F \subset K$ is called a *finite extension* if its degree $[K : F]$ is finite. Extensions of degree 2 are also called *quadratic* extensions, those of degree 3 are called *cubic* extensions, and so on. The degree of an extension $F \subset K$ is 1 if and only if $F = K$.

The term *degree* comes from the case that $K = F(\alpha)$ is generated by one algebraic element α. In that case, K has the basis $(1, \alpha, \ldots, \alpha^{n-1})$, where n is the degree of the irreducible polynomial for α over F [Proposition (2.7)]. Thus we find the first important property of the degree:

(3.2) **Proposition.** If α is algebraic over F, then $[F(\alpha) : F]$ is the degree of the irreducible polynomial for α over F. □

This degree is also called the *degree of* α *over* F. Note that an element α has degree 1 over F if and only if it is an element of F, and α has degree ∞ if and only if it is transcendental over F.

Extensions of degree 2 are easy to describe.

(3.3) **Proposition.** Assume that the field F does not have characteristic 2, that is, that $1 + 1 \neq 0$ in F. Then any extension $F \subset K$ of degree 2 can be obtained by adjoining a square root: $K = F(\delta)$, where $\delta^2 = D$ is an element of F. Conversely, if δ is an element of an extension of F, and if $\delta^2 \in F$ but $\delta \notin F$, then $F(\delta)$ is a quadratic extension.

Proof. We first show that every quadratic extension is obtained by adjoining a root of a quadratic polynomial $f(x) \in F[x]$. To do this, we choose any element α of K which is not in F. Then $(1, \alpha)$ is a linearly independent set over F. Since K has dimension 2 as a vector space over F, $(1, \alpha)$ is a basis for K over F, and $K = F[\alpha]$. It follows that α^2 is a linear combination of $(1, \alpha)$, say $\alpha^2 = -b\alpha - c$, with $b, c \in F$. Then α is a root of $f(x) = x^2 + bx + c$.

Since $2 \neq 0$ in F, we can use the quadratic formula $\alpha = \frac{1}{2}(-b + \sqrt{b^2-4c})$ to solve the equation $x^2 + bx + c = 0$. This is proved by direct calculation. There are two choices for the square root, one of which gives our chosen root α. Let δ denote that choice: $\delta = \sqrt{b^2-4c} = 2\alpha + b$. Then δ is in K, and it also generates K over F. Its square is the discriminant $b^2 - 4c$, which is in F.

The last assertion of the proposition is clear. □

The second important property of the degree is that it is multiplicative in towers of fields.

(3.4) **Theorem.** Let $F \subset K \subset L$ be fields. Then $[L : F] = [L : K][K : F]$.

Proof. Let $\mathbf{B} = (y_1, \ldots, y_n)$ be a basis for L as a K-vector space, and let $\mathbf{C} = (x_1, \ldots, x_m)$ be a basis for K as an F-vector space. So $[L : K] = n$ and $[K : F] = m$. We will show that the set of mn products $\mathbf{P} = (\ldots, x_i y_j, \ldots)$ is a basis of L as an F-vector space, and this will prove the proposition. The same reasoning will work if $\mathbf{B}$ or $\mathbf{C}$ is infinite.

Let α be an element of L. Since $\mathbf{B}$ is a basis for L over K, we can write $\alpha = \beta_1 y_1 + \cdots + \beta_n y_n$, with $\beta_j \in K$, in a unique way. Since $\mathbf{C}$ is a basis for K over F, each β_i can be expressed uniquely, as $\beta_j = a_{1j} x_1 + \cdots + a_{mj} x_m$, with $a_{ij} \in F$. Thus $\alpha = \Sigma_{i,j} a_{ij} x_i y_j$. This shows that $\mathbf{P}$ spans L as an F-vector space. We know that β_j is uniquely determined by α, and since $\mathbf{B}$ is a basis for K over F, the elements a_{ij} are uniquely determined by β_j. So they are uniquely determined by α. This shows that $\mathbf{P}$ is linearly independent, and hence that it is a basis for L over F. □

One important case of a tower of field extensions is that K is a given extension of F and α is an element of K. Then the field $F(\alpha)$ generated by α is an intermediate field:

$$F \subset F(\alpha) \subset K. \tag{3.5}$$

(3.6) **Corollary.** Let K be an extension of F, of finite degree n. Let α be an element of K. Then α is algebraic over F, and its degree divides n.

To see this, we apply Theorem (3.4) to the fields $F \subset F(\alpha) \subset K$ and use the fact that the degree of α over F is $[F(\alpha) : F]$ if α is algebraic, while $[F(\alpha) : F] = \infty$ if α is transcendental. □

Here are some sample applications:

(3.7) **Corollary.** Let K be a field extension of F of prime degree p, and let α be an element of K which is not in F. Then α has degree p over F, and $K = F(\alpha)$.

For, $p = [K : F] = [K : F(\alpha)][F(\alpha) : F]$. One of the terms on the right side is 1. Since $\alpha \notin F$, it is not the second term, so $[K : F(\alpha)] = 1$ and $[F(\alpha) : F] = p$. Therefore $K = F(\alpha)$. □

(3.8) **Corollary.** Every irreducible polynomial in $\mathbb{R}[x]$ has degree 1 or 2.

We proved this in Chapter 11, Section 1, but let us derive it once more: Let g be an irreducible real polynomial. Then g has a root α in $\mathbb{C}$. Since $[\mathbb{C} : \mathbb{R}] = 2$, the degree of α over $\mathbb{R}$ divides 2, by (3.6). Therefore the degree of g is 1 or 2. □

(3.9) **Examples.**

(a) Let $\alpha = \sqrt[3]{2}$, $\beta = \sqrt[4]{5}$. Consider the field $L = \mathbb{Q}(\alpha, \beta)$ obtained by adjoining α and β to $\mathbb{Q}$. Then $[L : \mathbb{Q}] = 12$. For L contains the subfield $\mathbb{Q}(\alpha)$, which has degree 3 over $\mathbb{Q}$, because the irreducible polynomial for α over $\mathbb{Q}$ is $x^3 - 2$. Therefore 3 divides $[L : \mathbb{Q}]$. Similarly, L contains $\mathbb{Q}(\beta)$ and β has de-

gree 4 over $\mathbb{Q}$, so 4 divides $[L : \mathbb{Q}]$. On the other hand, the degree of β over the field $\mathbb{Q}(\alpha)$ is at most 4, because β is a root of $x^4 - 5$, and this polynomial has coefficients in $\mathbb{Q}(\alpha)$. The chain of fields $L = \mathbb{Q}(\alpha,\beta) \supset \mathbb{Q}(\alpha) \supset \mathbb{Q}$ shows that $[L : \mathbb{Q}]$ is at most 12. So $[L : \mathbb{Q}] = 12$.

(b) It follows by reducing modulo 2 that the polynomial $f(x) = x^4 + 2x^3 + 6x^2 + x + 9$ is irreducible over $\mathbb{Q}$ [Chapter 11 (4.3)]. Let γ be a root of $f(x)$. Then there is no way to express $\alpha = \sqrt[3]{2}$ rationally in terms of γ, that is, $\alpha \notin \mathbb{Q}(\gamma)$. For $[\mathbb{Q}(\alpha) : \mathbb{Q}] = 3$, $[\mathbb{Q}(\gamma) : \mathbb{Q}] = 4$, and 3 does not divide 4. So we can't have $\mathbb{Q}(\gamma) > \mathbb{Q}(\alpha)$. On the other hand, since i has degree 2 over $\mathbb{Q}$, it is not so easy to decide whether i is in $\mathbb{Q}(\gamma)$. (In fact, it is not.) □

The next two theorems state the most important abstract consequences of the multiplicative property of degrees.

(3.10) **Theorem.** Let K be an extension of F. The elements of K which are algebraic over F form a subfield of K.

Proof. Let α,β be algebraic elements of K. We must show that $\alpha + \beta$, $\alpha\beta$, $-\alpha$, and α^{-1} (if $\alpha \neq 0$) are algebraic too. We note that since α is algebraic, $[F(\alpha) : F] < \infty$. Moreover, β is algebraic over F, and hence it is also algebraic over the bigger field $F(\alpha)$. Therefore the field $F(\alpha,\beta)$, which is generated over $F(\alpha)$ by β, is a finite extension of $F(\alpha)$, that is, $[F(\alpha,\beta) : F(\alpha)] < \infty$. By Theorem (3.4), $[F(\alpha,\beta) : F]$ is finite too. Therefore every element of $F(\alpha,\beta)$ is algebraic over F (3.6). The elements $\alpha + \beta$, $\alpha\beta$, etc. all lie in $F(\alpha,\beta)$, so they are algebraic. This proves that the algebraic elements form a field. □

Suppose for example that $\alpha = \sqrt{a}$, $\beta = \sqrt{b}$, where $a, b \in F$. Let us determine a polynomial having $\gamma = \alpha + \beta$ as a root. To do this, we compute the powers of γ, and we use the relations $\alpha^2 = a$, $\beta^2 = b$ to simplify when possible. Then we look for a linear relation among the powers:

$$\gamma^2 = \alpha^2 + 2\alpha\beta + \beta^2 = (a+b) + 2\alpha\beta$$

$$\gamma^4 = (a+b)^2 + 4(a+b)\alpha\beta + 4\alpha^2\beta^2 = (a^2+6ab+b^2) + 4(a+b)\alpha\beta.$$

We won't need the other powers because we can eliminate $\alpha\beta$ from these two equations to obtain the equation $\gamma^4 - 2(a+b)\gamma^2 + (a-b)^2 = 0$. Thus γ is a root of the polynomial

$$g(x) = x^4 - 2(a+b)x^2 + (a-b)^2,$$

which has coefficients in F, as required.

This method of undetermined coefficients will always produce a polynomial having an element such as $\alpha + \beta$ as a root, if the irreducible polynomials for α and β are known. Suppose that the degrees of two elements α,β are d_1, d_2, and let $n = d_1 d_2$. Any element of $F(\alpha,\beta)$ is a linear combination, with coefficients in F, of the n monomials $\alpha^i\beta^j$, $0 \leq i < d_1$, $0 \leq j < d_2$. This is because $F(\alpha,\beta) = F[\alpha,\beta]$ (2.6), and these monomials span $F[\alpha,\beta]$. Given an element $\gamma \in F(\alpha,\beta)$,

we write the powers $1, \gamma, \gamma^2, \ldots, \gamma^n$ as linear combinations of these monomials, with coefficients in F. Since there an $n + 1$ of the powers γ^ν and only n monomials $\alpha^i\beta^j$, the powers are linearly dependent. A linear dependence relation determines a polynomial with coefficients in F of which γ is a root.

But there is one point which complicates matters. Let $g(x)$ be the polynomial having γ as a root which we find in this way. This polynomial may be reducible. For instance, it may happen that γ is actually in the field F, though α, β aren't in F. If so, the method we described is unlikely to produce its irreducible equation $x - \gamma$. It is harder to determine the *irreducible* polynomial for γ over F. □

An extension K of a field F is called an *algebraic extension*, and K is said to be *algebraic over* F, if all its elements are algebraic.

(3.11) **Theorem.** Let $F \subset K \subset L$ be fields. If L is algebraic over K and K is algebraic over F, then L is algebraic over F.

Proof. We need to show that every element $\alpha \in L$ is algebraic over F. We are given that α is algebraic over K, hence that some equation of the form

$$\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_1\alpha + a_0 = 0$$

holds, with $a_0, \ldots, a_{n-1} \in K$. Therefore α is algebraic over the field $F(a_0, \ldots, a_{n-1})$ generated by $a_0, \ldots, a_{n-1}$ over F. Note that each coefficient a_i, being in K, is algebraic over F. We consider the chain of fields

$$F \subset F(a_0) \subset F(a_0, a_1) \subset \cdots \subset F(a_0, a_1, \ldots, a_{n-1}) \subset F(a_0, a_1, \ldots, a_{n-1}, \alpha)$$

obtained by adjoining the elements $a_0, \ldots, a_{n-1}, \alpha$ in succession. For each i, a_{i+1} is algebraic over $F(a_0, \ldots, a_i)$ because it is algebraic over F. Also, α is algebraic over $F(a_0, a_1, \ldots, a_{n-1})$. So each extension in the chain is finite. By Theorem (3.4), the degree of $F(a_0, a_1, \ldots, a_{n-1}, \alpha)$ over F is finite. Therefore by Corollary (3.6) α is algebraic over F. □

4. CONSTRUCTIONS WITH RULER AND COMPASS

There are famous theorems which assert that certain geometric constructions, such as trisection of an angle, can not be done with ruler and compass alone. We will now use the concept of degree of a field extension to prove some of them.

Here are the rules for basic ruler and compass construction:

(4.1)

(a) Two points in the plane are given to start with. These points are considered to be *constructed*.

(b) If two points have been constructed, we may draw the line through them, or draw a circle with center at one point and passing through the other. Such lines and circles are then considered to be *constructed*.

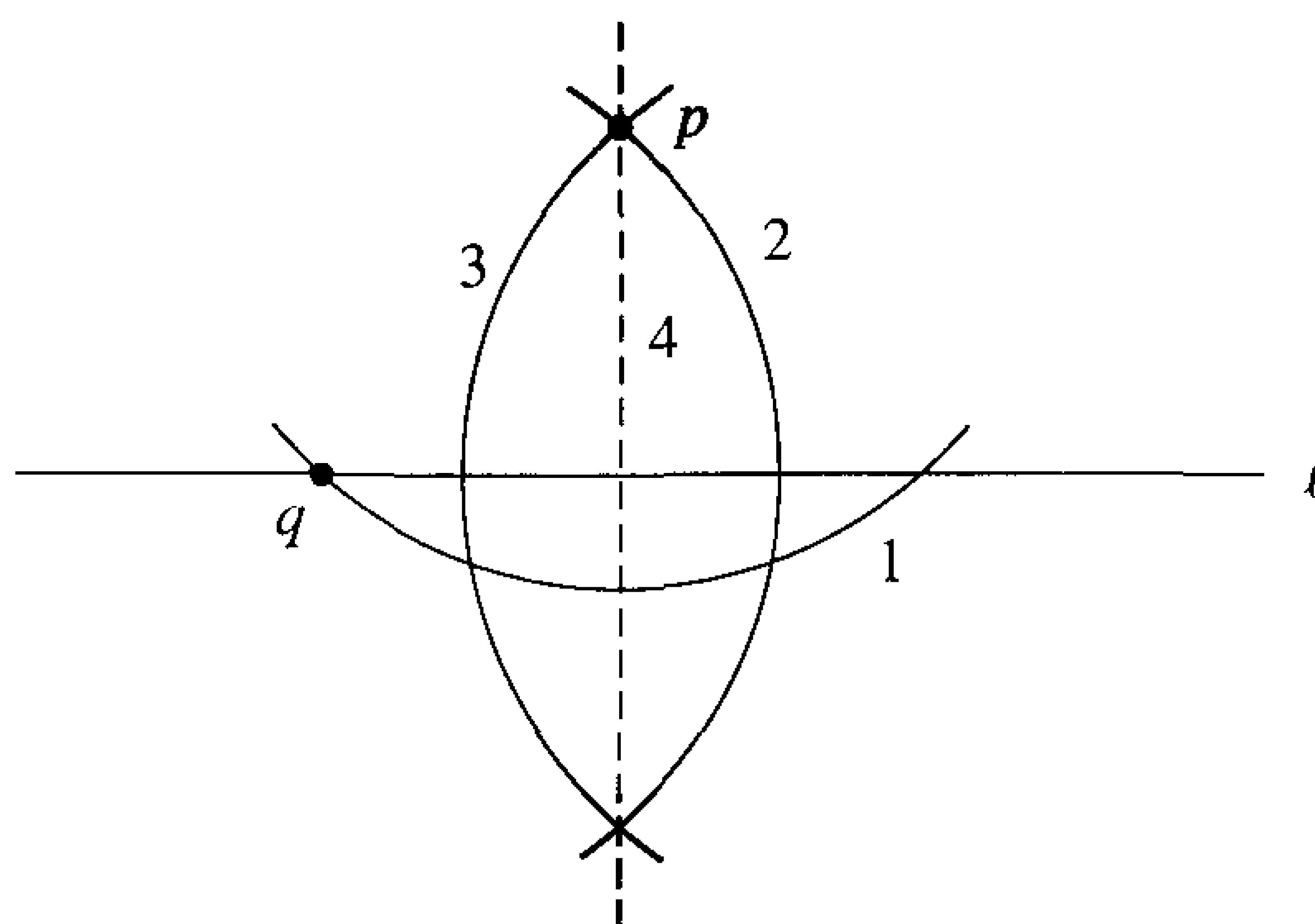
(c) The points of intersection of lines and circles which have been constructed are considered to be *constructed.*

Note that our ruler may be used only to draw straight lines through constructed points. We are not allowed to use it for measurement. Sometimes it is referred to as a "straight-edge" to make this point clear.

We will describe all possible constructions, beginning with some familiar ones. In each figure, the lines and circles are to be drawn in the order indicated.

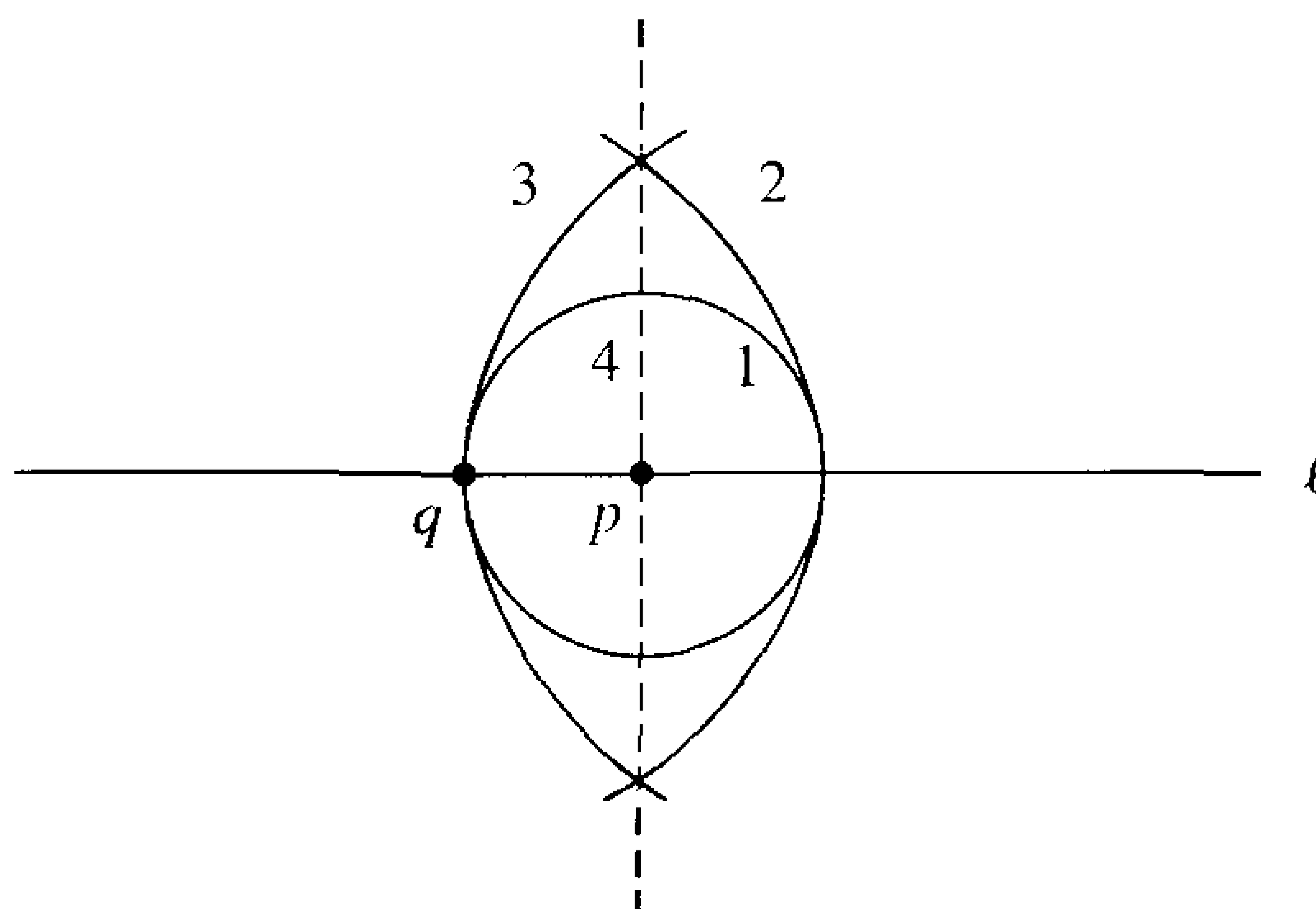
(4.2) **Construction.** Draw a line through a constructed point p and perpendicular to a constructed line ℓ.

Case 1: $p \notin \ell$

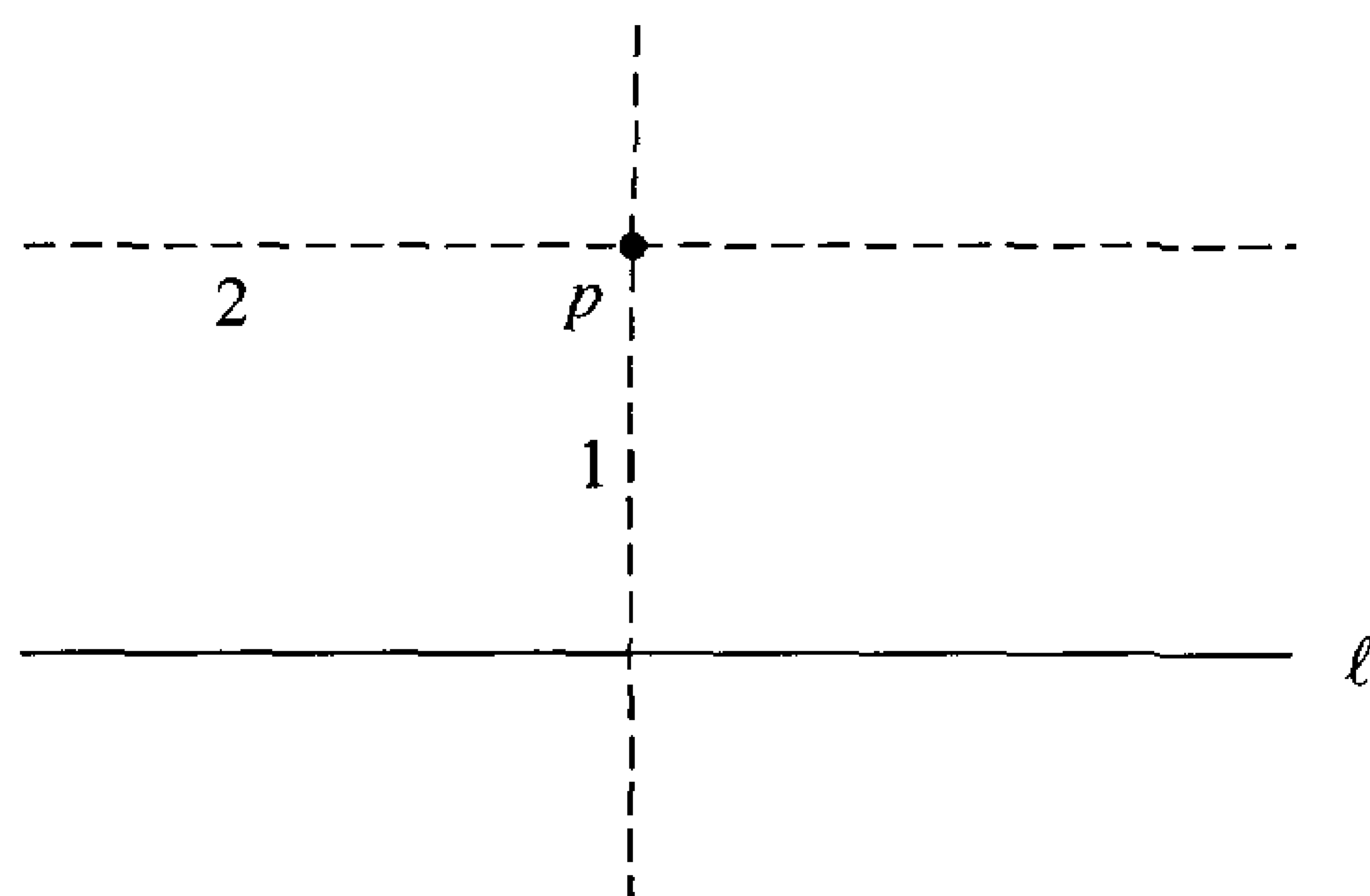


This construction works with any point $q \in \ell$ which is not on the perpendicular. However, we had better not choose points arbitrarily, because if we do we'll have difficulty keeping track of which points we have constructed and which ones are merely artifacts of an arbitrary choice. Whenever we want an arbitrary point, we will construct a particular one for the purpose.

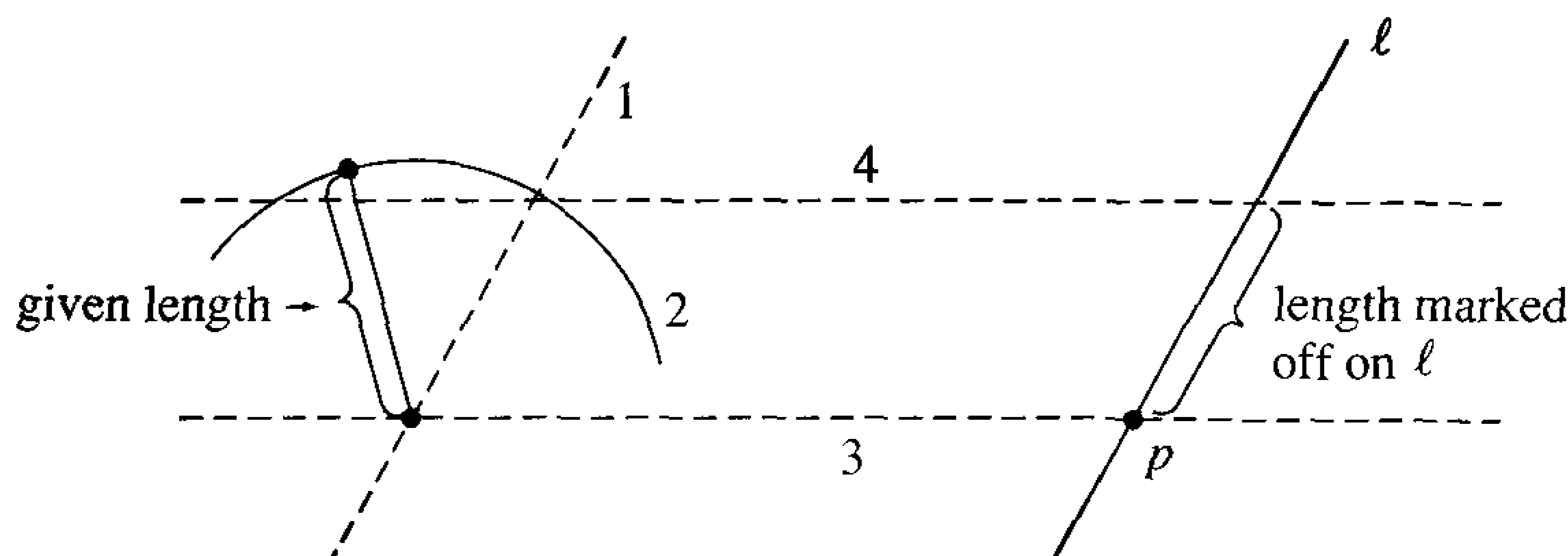
Case 2: $p \in \ell$



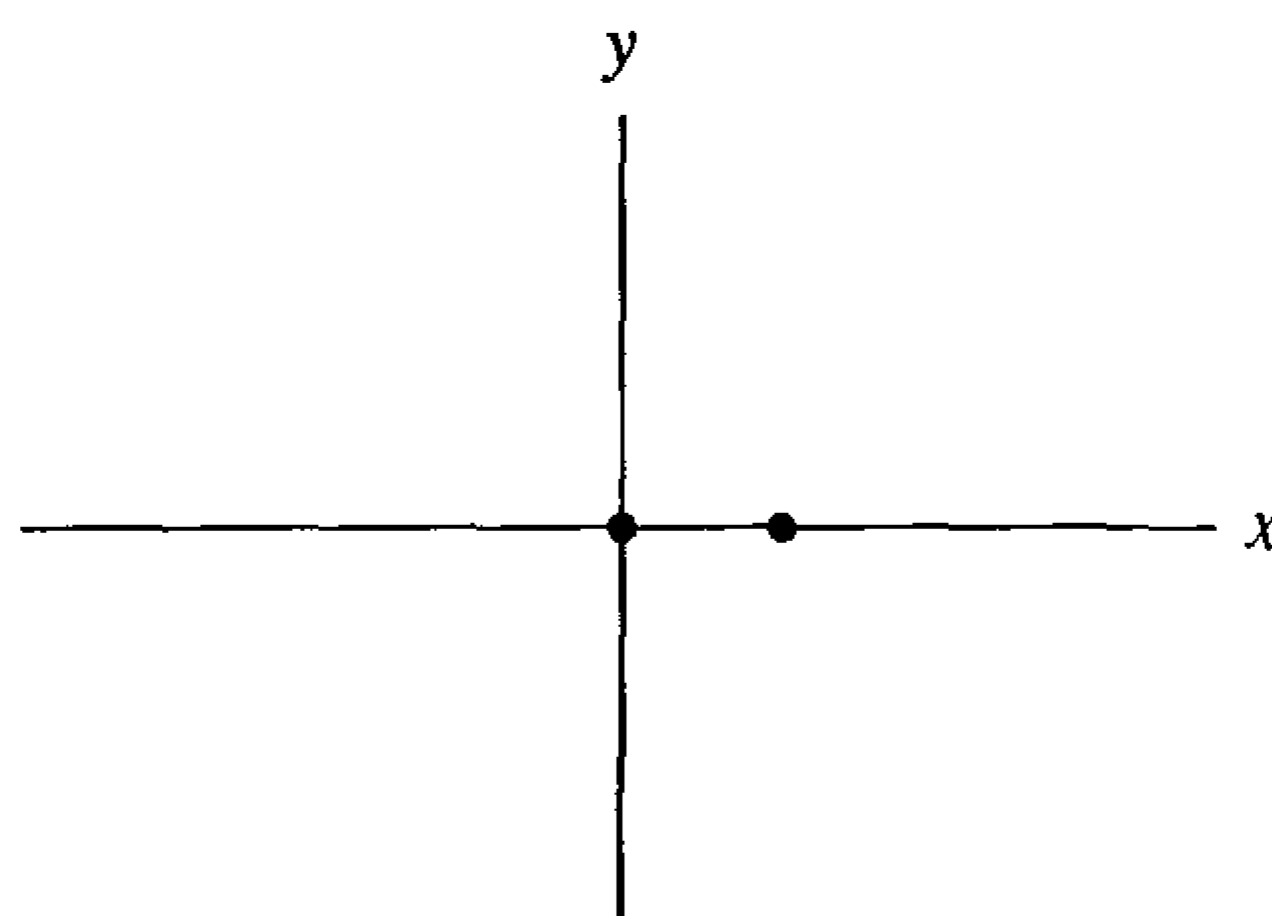
(4.3) **Construction.** Draw a line parallel to ℓ and passing through p. Apply Cases 1 and 2 above:



(4.4) **Construction.** Mark off a length defined by two points onto a constructed line ℓ, starting at a constructed point $p \in \ell$. Use construction of parallels.



These constructions allow us to introduce Cartesian coordinates into the plane so that the two points which are given to us to start have coordinates $(0, 0)$ and $(0, 1)$. Other choices of coordinate systems could be used, but they lead to equivalent theories.



We will call a real number a *constructible* if its absolute value $|a|$ is the distance between two constructible points, the unit length being the distance between the points given originally.

(4.5) **Proposition.** A point $p = (a, b)$ is constructible if and only if its Cartesian coordinates a and b are constructible numbers.

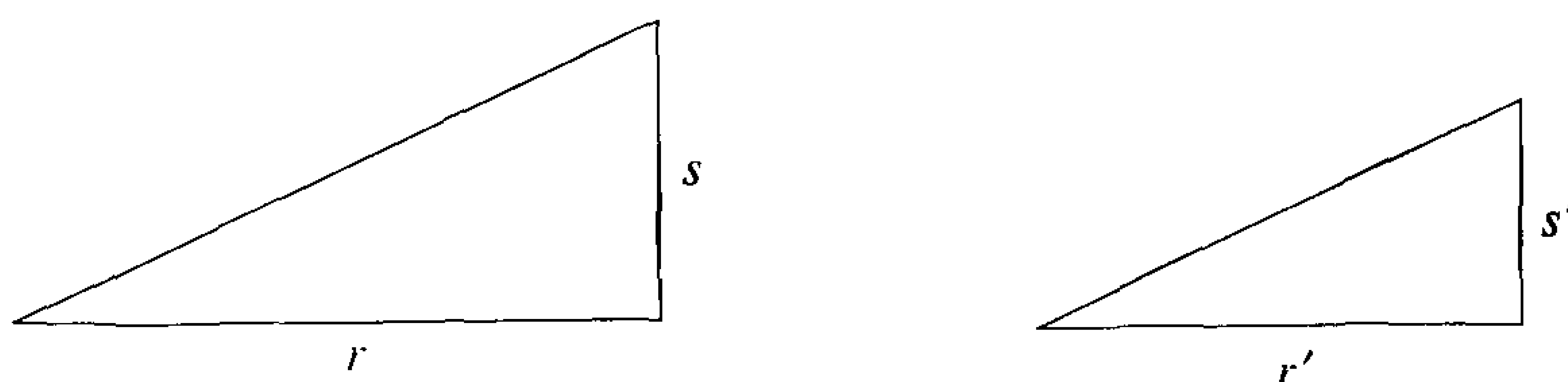
Proof. This follows from the above constructions. Given a point p, we can construct its coordinates by dropping perpendiculars to the axes. Conversely, if a and b are given constructible numbers, then we can construct the point p by marking a, b off on the two axes using (4.4) and erecting perpendiculars. □

(4.6) **Proposition.** The constructible numbers form a subfield of $\mathbb{R}$.

Proof. We will show that if a and b are positive constructible numbers, then $a + b$, ab, $a - b$, (if $a > b$), and a^{-1} (if $a \neq 0$) are also constructible. The closure in case a or b is negative follows easily.

Addition and subtraction are done by marking lengths on a line, using Construction (4.4).

For multiplication, we use similar right triangles:



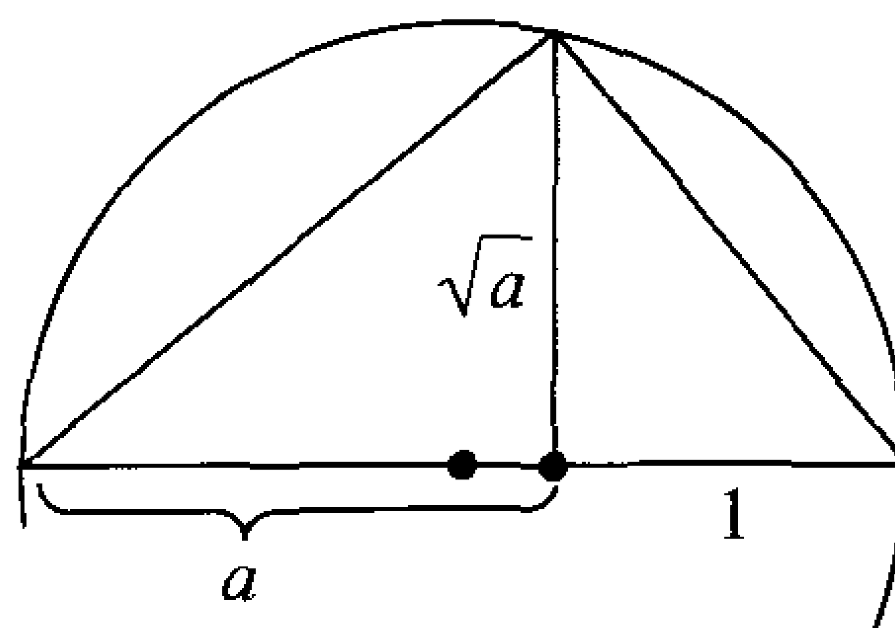
Given one triangle and one side of a second triangle, the second triangle can be constructed by parallels.

To construct the product ab, we take $r = 1$, $s = a$, and $r' = b$. Then since $r/s = r'/s'$, it follows that $s' = ab$. To construct a^{-1}, we take $r = a$, $s = 1$, and $r' = 1$. Then $s' = a^{-1}$. □

(4.7) **Proposition.** If a is a positive constructible number, then so is $\sqrt{a}$.

Proof. We use similar triangles again. We must construct them so that $r = a$, $r' = s$, and $s' = 1$. Then $s = r' = \sqrt{a}$.

How to make the construction is less obvious this time, but we can use inscribed triangles in a circle. A triangle inscribed into a circle, with a diameter as its hypotenuse, is a right triangle. This is a theorem of high school geometry. It can be checked using the equation for a circle and Pythagoras's theorem. So we draw a circle whose diameter is $1 + a$ and proceed as in the figure below. Note that the large triangle is divided into two similar triangles.



(4.8) **Proposition.** Suppose four points are given, whose coordinates are in a subfield F of $\mathbb{R}$. Let A, B be lines or circles drawn using the given points. Then the

points of intersection of A and B have coordinates of F, or in a field of the form $F(\sqrt{r})$, where r is a positive number in F.

Proof. The line through (a_0, b_0), (a_1, b_1) has the linear equation

$$(a_1 - a_0)(y - b_0) = (b_1 - b_0)(x - a_0).$$

The circle with center (a_0, b_0) and passing through (a_1, b_1) has the quadratic equation

$$(x - a_0)^2 + (y - b_0)^2 = (a_1 - a_0)^2 + (b_1 - b_0)^2.$$

The intersection of two lines can be found by solving two linear equations whose coefficients are in F. So its coordinates are in F too. To find the intersection of a line and a circle, we use the equation of the line to eliminate one variable from the equation of the circle, obtaining a quadratic equation in one unknown. This quadratic equation has solutions in the field $F(\sqrt{D})$, where D is the discriminant, which is an element of F. If $D < 0$, the line and circle do not intersect.

Consider the intersection of two circles, say

$$(x - a_1)^2 + (y - b_1)^2 = r_1^2 \quad \text{and} \quad (x - a_2)^2 + (y - b_2)^2 = r_2^2,$$

where $a_i, b_i, r_i \in F$. In general, the solution of a pair of quadratic equations in two variables requires solving an equation of degree 4. In this case we are lucky: The difference of the two quadratic equations is a linear equation which we can use to eliminate one variable, as before. □

(4.9) **Theorem.** Let $a_1, \ldots, a_m$ be constructible real numbers. There is a chain of subfields $\mathbb{Q} = F_0 \subset F_1 \subset F_2 \subset \cdots \subset F_n = K$ such that

(i) K is a subfield of $\mathbb{R}$;
(ii) $a_1, \ldots, a_m \in K$;
(iii) for each $i = 0, \ldots, n-1$, the field F_{i+1} is obtained from F_i by adjoining the square root of a positive number $r_i \in F_i$, which is not a square in F_i.

Conversely, let $\mathbb{Q} = F_0 \subset F_1 \subset \cdots \subset F_n = K$ be a chain of subfields of $\mathbb{R}$ which satisfies (iii) . Then every element of K is constructible.

Proof. We introduced coordinates so that the points originally given have coordinates in $\mathbb{Q}$. The process of constructing the numbers a_i involves drawing lines and circles and taking their intersections. So the first assertion follows by induction from Proposition (4.8). Conversely, if such a tower of fields is given, then its elements are constructible, by Propositions (4.6) and (4.7). □

(4.10) **Corollary.** If a is a constructible real number, then it is algebraic, and its degree over $\mathbb{Q}$ is a power of 2.

For, in the chain of fields (4.9), the degree of F_{i+1} over F_i is 2, and hence $[K : \mathbb{Q}] = 2^n$. Corollary (3.6) tells us that the degree of a divides 2^n, hence that it is a power of 2. □

The converse of Corollary (4.10) is false. There exist real numbers a which have degree 4 over $\mathbb{Q}$ but which are not constructible. We will be able to prove this later, using Galois theory.

We can now prove the impossibility of certain geometric constructions. Our method will be to show that if a certain construction were possible, then it would also be possible to construct an algebraic number whose degree over $\mathbb{Q}$ is not a power of 2. This would contradict (4.10).

Let us discuss trisection of the angle as the first example. We must pose the problem carefully, because many angles, 45° for instance, can be trisected. The customary way to state the problem is to ask for a single method of construction which will work for *any given angle*.

To be as specific as possible, let us say that an angle θ is *constructible* if its cosine $\cos\theta$ is constructible. Other equivalent definitions are possible. For example, with this definition, θ is constructible if and only if the line which passes through the origin and meets the x-axis in the angle θ is constructible. Or, θ is constructible if and only if it is possible to construct any two lines meeting in an angle θ.

Now just giving an angle θ (say by marking off its cosine on the x-axis) provides us with new information which may be used in a hypothetical trisection. To analyze the consequences of this new information, we should start over and determine all constructions which can be made when, in addition to two points, one more length ($= \cos\theta$) is given at the start. We would rather not take the time to do this, and there is a way out. We will exhibit a particular angle θ with these properties:

(4.11) (i) θ is constructible, and
(ii) $\frac{1}{3}\theta$ is not constructible.

The first condition tells us that being given the angle θ provides no new information for us: If the angle θ can be trisected when given, it can also be trisected without being given. The second condition tells us that there is no general method of trisection, because there is no way to trisect θ.

The angle $\theta = 60°$ does the job. A 60° angle is constructible because $\cos 60° = \frac{1}{2}$. On the other hand, it is impossible to construct a 20° angle. To show this, we will show that $\cos 20°$ is an algebraic number of degree 3 over $\mathbb{Q}$. Then Corollary (4.10) will show that $\cos 20°$ is not constructible, hence that 60° can not be trisected.

The addition formulas for sine and cosine can be used to prove the identity

$$\cos 3\theta = 4\cos^3\theta - 3\cos\theta. \tag{4.12}$$

Setting $\theta = 20°$ and $\alpha = \cos 20°$, we obtain the equation $\frac{1}{2} = 4\alpha^3 - 3\alpha$, or $8\alpha^3 - 6\alpha - 1 = 0$.

(4.13) **Lemma.** The polynomial $f(x) = 8x^3 - 6x - 1$ is irreducible over $\mathbb{Q}$.

Proof. It is enough to check for linear factors $ax + b$, where a, b are integers such that a divides 8, and $b = \pm 1$. Another way to prove irreducibility is to check that f has no root modulo 5. □

This lemma tells us that α has degree 3 over $\mathbb{Q}$, hence that it can not be constructed.

As another example, let us show that the regular 7-gon can not be constructed. This is similar to the above problem: The construction of 20° is equivalent to the construction of the 18-gon. Let θ denote the angle $2\pi/7$ and let $\zeta = \cos\theta + i\sin\theta$. Then ζ is a root of the equation $x^6 + x^5 + \cdots + 1 = 0$, which is irreducible [Chapter 11 (4.6)]. Hence ζ has degree 6 over $\mathbb{Q}$. If the 7-gon were constructible, then $\cos\theta$ and $\sin\theta$ would be constructible numbers, and hence they would lie in a real field extension of degree 2^n over $\mathbb{Q}$, by Theorem (4.9). Call this field K, and consider the extension $K(i)$. This extension has degree 2. Therefore $[K(i) : \mathbb{Q}] = 2^{n+1}$. But $\zeta = \cos\theta + i\sin\theta \in K(i)$. This contradicts the fact that the degree of ζ is 6 (3.6).

Notice that this argument is not special to the number 7. It applies to any prime integer p, provided only that $p - 1$, the degree of the irreducible polynomial $x^{p-1} + \cdots + x + 1$, is not a power of 2.

(4.14) **Corollary.** Let p be a prime integer. If the regular p-gon can be constructed by ruler and compass, then $p = 2^r + 1$ for some integer r. □

Gauss proved the converse: If a prime has the form $2^r + 1$, then the regular p-gon can be constructed. The regular 17-gon, for example, can be constructed with ruler and compass. We will learn how to prove this in the next chapter.

5. SYMBOLIC ADJUNCTION OF ROOTS

Up to this point, we have used subfields of the complex numbers as our examples. Abstract constructions are not needed to create these fields (except that the construction of $\mathbb{C}$ from $\mathbb{R}$ is abstract). We simply adjoin complex numbers to the rational numbers as desired and work with the subfield they generate. But finite fields and function fields are not subfields of a familiar, all-encompassing field analogous to $\mathbb{C}$, so these fields must be constructed. The fundamental tool for their construction is the adjunction of elements to a ring, which we studied in Section 5 of Chapter 10. It is applied here to the case that the ring we start with is a field F.

Let us review this construction. Given a polynomial $f(x)$ with coefficients in F, we may adjoin an element α satisfying the polynomial equation $f(\alpha) = 0$ to F. The abstract procedure is to form the polynomial ring $F[x]$ and then take the quotient ring

$$R' = F[x]/(f). \tag{5.1}$$

This construction always yields a ring R' and a homomorphism $F \longrightarrow R'$, such that the residue $\bar{x}$ of x satisfies the relation $f(\bar{x}) = 0$.

However, we want to construct not only a ring, but a field, and here the theory of polynomials over a field comes into play. Namely, that theory tells us that the principal ideal (f) is a maximal ideal if and only if f is irreducible [Chapter 11 (1.6)]. Therefore the ring R' will be a field if and only if f is an irreducible polynomial.

(5.2) **Lemma.** Let F be a field, and let f be an irreducible polynomial in $F[x]$. Then the ring $K = F[x]/(f)$ is an extension field of F, and the residue $\overline{x}$ of x is a root of $f(x)$ in K.

Proof. The ring K is a field because (f) is a maximal ideal. Also, the homomorphism $F \longrightarrow K$, which sends the elements of F to the residues of the constant polynomials, is injective, because F is a field. So we may identify F with its image, a subfield of K. The field K becomes an extension of F by means of this identification. Finally, $\overline{x}$ satisfies the equation $f(\overline{x}) = 0$, which means that it is a root of f. □

(5.3) **Proposition.** Let F be a field, and let $f(x)$ be a monic polynomial in $F[x]$ of positive degree. There exists a field extension K of F such that $f(x)$ factors into linear factors over K.

Proof. We use induction on the degree of f. The first case is that f has a root α in F, so that $f(x) = (x - \alpha)g(x)$ for some polynomial g. If so, we replace f by g, and we are done by induction. Otherwise, we choose an irreducible factor $g(x)$ of $f(x)$. By Lemma (5.2), there is a field extension of F, call it F_1, in which $g(x)$ has a root α. We replace F by F_1 and are thereby reduced to the first case. □

As we have seen, the polynomial ring $F[x]$ is an important tool for studying extensions of a field F. When we are working with two fields at the same time, there is an interplay between their polynomial rings. This interplay doesn't present serious difficulties, but instead of scattering the points which need to be mentioned about in the text, we have collected them here.

Notice that if K is an extension field of F, then the polynomial ring $K[x]$ contains $F[x]$ as subring. So computations which are made in the ring $F[x]$ are also valid in $K[x]$.

(5.4) **Proposition.** Let f and g be polynomials with coefficients in a field F, and let K be an extension field of F.

(a) Division with remainder of g by f gives the same answer, whether carried out in $F[x]$ or in $K[x]$.
(b) f divides g in $K[x]$ if and only if f divides g in $F[x]$.
(c) The monic greatest common divisor d of f and g is the same, whether computed in $F[x]$ or in $K[x]$.
(d) If f and g have a common root in K, then they are not relatively prime in $F[x]$. Conversely, if f and g are not relatively prime in $F[x]$, then there exists an extension field L in which they have a common root.
(e) If f is irreducible in $F[x]$ and if f and g have a common root in K, then f divides g in $F[x]$.

Proof. (a) Carry out the division in $F[x] : g = fq + r$. This equation also holds in the bigger ring $K[x]$, and further division of the remainder by f is not possible, because r has lower degree than f, or else it is zero.

(b) This is the case that the remainder is zero in (a).

(c) Let d, d' denote the monic greatest common divisors of f and g in $F[x]$ and in $K[x]$. Then d is also a common divisor in $K[x]$. So d divides d' in $K[x]$, by definition of d'. In addition, we know that d has the form $d = pf + qg$, for some elements $p, q \in F[x]$. Since d' divides f and g, it divides $pf + qg = d$ too. Thus d and d' are associates in $K[x]$, and, being monic, they are equal.

(d) Let α be a common root of f and g in K. Then $x - \alpha$ is a common divisor of f and g in $K[x]$. So their greatest common divisor in $K[x]$ is not 1. By (c), it is not 1 in $F[x]$ either. Conversely, if f and g have a common divisor d of degree > 0, then by (5.3), d has a root in some extension field L. This root will be a common root of f and g.

(e) If f is irreducible, then its only divisors in $F[x]$ are $1, f$, and their associates. Part (d) tells us that the greatest common divisor of f and g in $F[x]$ is not 1. Therefore it is f. □

The final topic of this section concerns the derivative $f'(x)$ of a polynomial $f(x)$. In algebra, the derivative is computed using the rules from calculus for differentiating polynomial functions. In other words, we define the derivative of x^n to be the polynomial nx^{n-1}, and if $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$, then

$$f'(x) = na_nx^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1. \tag{5.5}$$

The integer coefficients in this formula are to be interpreted as elements of F by means of the homomorphism $\mathbb{Z} \longrightarrow F$ [Chapter 10 (3.18)]. So the derivative is a polynomial with coefficients in the same field. It can be shown that rules such as the product rule for differentiation hold.

Though differentiation is an algebraic procedure, there is no a priori reason to suppose that it has much algebraic significance; however, it does. For us, the most important property of the derivative is that it can be used to recognize multiple roots of a polynomial.

(5.6) **Lemma.** Let F be a field, let $f(x) \in F[x]$ be a polynomial, and let $\alpha \in F$ be a root of $f(x)$. Then α is a *multiple root*, meaning that $(x - \alpha)^2$ divides $f(x)$, if and only if it is a root of both $f(x)$ and $f'(x)$.

Proof. If α is a root of f, then $x - \alpha$ divides f: $f(x) = (x - \alpha)g(x)$. Then α is a root of g if and only if it is a multiple root of f. By the product rule for differentiation,

$$f'(x) = (x - \alpha)g'(x) + g(x).$$

Substituting $x = \alpha$ shows that $f'(\alpha) = 0$ if and only if $g(\alpha) = 0$. □

(5.7) **Proposition.** Let $f(x) \in F[x]$ be a polynomial. There exists a field extension K of F in which f has a multiple root if and only if f and f' are not relatively prime.

Proof. If f has a multiple root in K, then f and f' have a common root in K by Lemma (5.6), and so they are not relatively prime in K. Hence they are not relatively prime in F either. Conversely, if f and f' are not relatively prime, then they have a common root in some field extension K, hence f has a multiple root there. □

Here is one of the most important applications of the derivative to field theory:

(5.8) **Proposition.** Let f be an irreducible polynomial in $F[x]$. Then f has no multiple root in any field extension of F unless the derivative f' is the zero polynomial. In particular, if F is a field of characteristic zero, then f has no multiple root.

Proof. By the previous proposition, we must show that f and f' are relatively prime unless f' is the zero polynomial. Since f is irreducible, the only way that it can have a nonconstant factor in common with another polynomial g is for f to divide g (5.4e). And if f divides g, then $\deg g \geq \deg f$, or else $g = 0$. Now the degree of the derivative f' is less than the degree of f. So f and f' have no nonconstant factor in common unless $f' = 0$, as required. In a field of characteristic zero, the derivative of a nonconstant polynomial is not zero. □

The derivative of a nonconstant polynomial $f(x)$ may be identically zero if F has prime characteristic p. This happens when the exponent of every monomial occurring in f is divisible by p. A typical polynomial whose derivative is zero in characteristic 5 is

$$f(x) = x^{15} + ax^{10} + bx^5 + c,$$

where a, b, c can be arbitrary elements of F. Since the derivative of this polynomial is identically zero, its roots in any extension field are all multiple roots. Whether or not this polynomial is irreducible depends on F and on a, b, c.

6. FINITE FIELDS

In this section, we describe all fields having finitely many elements. We remarked in Section 1 that a finite field K contains one of the prime fields $\mathbb{F}_p$, and of course since K is finite, it will be finite-dimensional when considered as a vector space over this field. Let us denote $\mathbb{F}_p$ by F, and let r denote the degree $[K : F]$. As an F-vector space, K is isomorphic to the space F^r, and this space contains p^r elements. So the order of a finite field is always a power of a prime. It is customary to use the letter q for this number:

$$q = p^r = |K|. \tag{6.1}$$

When referring to finite fields, p will always denote a prime integer and q a power of p, the number of elements, or *order*, of the field.

Fields with q elements are often denoted by $\mathbb{F}_q$. We are going to show that all fields with the same number of elements are isomorphic, so this notation is not too ambiguous. However, the isomorphism will not be unique when $r > 1$.

The simplest example of a finite field other than the prime field $\mathbb{F}_p$ is the field $K = \mathbb{F}_4$ of order 4. There is a unique irreducible polynomial $f(x)$ of degree 2 in $\mathbb{F}_2[x]$, namely

$$f(x) = x^2 + x + 1 \tag{6.2}$$

[see Chapter 11 (4.3)], and the field K is obtained by adjoining a root α of $f(x)$ to $F = \mathbb{F}_2$:

$$K \approx F[x]/(x^2 + x + 1).$$

The order of this field is 4 because α has degree 2, which tells us that K has dimension 2 as a vector space over the field F.

The set $(1, \alpha)$ forms a basis of K over F, so the elements of K are the four linear combinations of these two elements, with mod-2 coefficients 0, 1. They are

$$\{0, 1, \alpha, 1 + \alpha\} = \mathbb{F}_4. \tag{6.3}$$

The element $1 + \alpha$ is the second root of the polynomial $f(x)$ in K. Computation in K is made using the relations $1 + 1 = 0$ and $\alpha^2 + \alpha + 1 = 0$. *Do not confuse the field* $\mathbb{F}_4$ *with the ring* $\mathbb{Z}/(4)$!

Here are the main facts about finite fields:

(6.4) **Theorem.** Let p be a prime, and let $q = p^r$ be a power of p, with $r \geq 1$.

(a) There exists a field of order q.
(b) Any two fields of order q are isomorphic.
(c) Let K be a field of order q. The multiplicative group $K^\times$ of nonzero elements of K is a cyclic group of order $q - 1$.
(d) The elements of K are roots of the polynomial $x^q - x$. This polynomial has distinct roots, and it factors into linear factors in K.
(e) Every irreducible polynomial of degree r in $\mathbb{F}_p[x]$ is a factor of $x^q - x$. The irreducible factors of $x^q - x$ in $\mathbb{F}_p[x]$ are precisely the irreducible polynomials in $\mathbb{F}_p[x]$ whose degree divides r.
(f) A field K of order q contains a subfield of order $q' = p^k$ if and only if k divides r.

The proof of this theorem is not very difficult, but since there are several parts, it will take some time. To motivate it, we will look at a few consequences first.

The striking aspect of (c) is that all nonzero elements of K can be listed as powers of a single suitably chosen one. This is not obvious, even for the prime field $\mathbb{F}_p$. For example, the residue of 3 is a generator of $\mathbb{F}_7^\times$. Its powers $3^0, 3^1 3^2, \ldots$ list the nonzero elements of $\mathbb{F}_7$ in the following order:

$$\mathbb{F}_7^\times = \{1, 3, 2, 6, 4, 5\}. \tag{6.5}$$

As another example, 2 is a generator of $\mathbb{F}_{11}^{\times}$, and its powers list that group in the order

$$(6.6) \qquad \mathbb{F}_{11}^{\times} = \{1, 2, 4, 8, 5, 10, 9, 7, 3, 6\}.$$

A generator for the cyclic group $\mathbb{F}_p^{\times}$ is called a *primitive element modulo p*. Note that the theorem does not tell us how to find a primitive element, only that one exists. Which residues modulo p are primitive elements is not well understood, but given a small prime p, we can find one by trial and error.

We now have two ways of listing the nonzero elements of $\mathbb{F}_p$, additively and multiplicatively:

$$(6.7) \qquad \mathbb{F}_p^{\times} = \{1, 2, 3, \ldots, p - 1\} = \{1, \nu, \nu^2, \ldots, \nu^{p-2}\},$$

where ν is a primitive element modulo p. Depending on the context, one or the other list may be the best for computation.

Of course, the additive group $\mathbb{F}_p^{+}$ of the prime field is always a cyclic group of order p. Both the additive and multiplicative structures of the prime field are very simple: They are cyclic. But the field structure of $\mathbb{F}_p$, governed by the distributive law, fits the two together in a subtle way.

Part (e) of the theorem is also striking. It is the basis for many methods of factoring polynomials modulo p. Let us look at a few cases in which q is a power of 2 as examples:

(6.8) **Examples.**

(a) The elements of the field $\mathbb{F}_4$ are the roots of the polynomial

$$(6.9) \qquad x^4 - x = x(x - 1)(x^2 + x + 1).$$

In this case, the irreducible factors of $x^4 - x$ in $\mathbb{Z}[x]$ happen to remain irreducible in $\mathbb{F}_2[x]$. Note that the factors of $x^2 - x$ appear here, because $\mathbb{F}_4$ contains $\mathbb{F}_2$.

Since we are working in characteristic 2, the signs are irrelevant: $x - 1 = x + 1$.

(b) The field $\mathbb{F}_8$ of order 8 has degree 3 over the prime field $\mathbb{F}_2$. Its elements are the eight roots of the polynomial

$$(6.10) \qquad x^8 - x = x(x - 1)(x^3 + x + 1)(x^3 + x^2 + 1), \quad \text{in } \mathbb{F}_2[x].$$

So the six elements in $\mathbb{F}_8$ which aren't in $\mathbb{F}_2$ fall into two classes: the three roots of $x^3 + x + 1$ and the three roots of $x^3 + x^2 + 1$.

The cubic factors of (6.10) are the two irreducible cubic polynomials of degree 3 in $\mathbb{F}_2[x]$ [see Chapter 11 (4.3)]. Notice that the irreducible factorization of this polynomial in the ring of integers is

$$(6.11) \qquad x^8 - x = x(x - 1)(x^6 + x^5 + \cdots + x + 1), \quad \text{in } \mathbb{Z}[x].$$

The third factor is reducible modulo 2.

To compute in the field $\mathbb{F}_8$, choose a root β of one of the cubics, say of $x^3 + x + 1$. Then $(1, \beta, \beta^2)$ is a basis of $\mathbb{F}_8$ as a vector space over $\mathbb{F}_2$. The elements

of $\mathbb{F}_8$ are the eight linear combinations with coefficients 0, 1:

$$\mathbb{F}_8 = \{0, 1, \beta, 1 + \beta, \beta^2, 1 + \beta^2, \beta + \beta^2, 1 + \beta + \beta^2\}. \tag{6.12}$$

Computation in $\mathbb{F}_8$ is done using the relation $\beta^3 + \beta + 1 = 0$.

Note that $\mathbb{F}_4$ is not contained in $\mathbb{F}_8$. It couldn't be, because $[\mathbb{F}_8 : \mathbb{F}_2] = 3$, $[\mathbb{F}_4 : \mathbb{F}_2] = 2$, and 2 does not divide 3.

(c) The field $\mathbb{F}_{16}$: The polynomial $x^{16} - x = x(x^{15} - 1)$ is divisible in $\mathbb{Z}[x]$ by $x^3 - 1$ and by $x^5 - 1$. Carrying out the division over the integers gives this factorization:

$$x^{16} - x = x(x - 1)(x^2 + x + 1)(x^4 + x^3 + x^2 + x + 1)(x^8 - x^7 + x^5 - x^4 + x^3 - x + 1). \tag{6.13}$$

This is the irreducible factorization in $\mathbb{Z}[x]$. But in $\mathbb{F}_2[x]$, the factor of degree 8 is not irreducible, and

$$x^{16} - x = x(x - 1)(x^2 + x + 1)(x^4 + x^3 + x^2 + x + 1)(x^4 + x^3 + 1)(x^4 + x + 1). \tag{6.14}$$

This factorization displays the three irreducible polynomials of degree 4 in $\mathbb{F}_2[x]$. Note that the factors of $x^4 - x$ appear among the factors of $x^{16} - x$. This agrees with the fact that $\mathbb{F}_{16}$ contains $\mathbb{F}_4$.

We will now begin the proof of Theorem (6.4). We will prove the various parts in the following order: (d), (c), (a), (b), (e), and (f).

Proof of Theorem (6.4d). Let K be a field of order q. The multiplicative group $K^\times$ has order $q - 1$. Therefore the order of any element $\alpha \in K^\times$ divides $q - 1$: $\alpha^{q-1} = 1$. This means that α is a root of the polynomial $x^{q-1} - 1$. The remaining element of K, zero, is a root of the polynomial x. So every element of K is a root of $x(x^{q-1} - 1) = x^q - x$. Since this polynomial has q distinct roots in K, it factors into linear factors in that field:

$$x^q - x = \prod_{\alpha \in K} (x - \alpha). \tag{6.15}$$

This proves part (d) of the theorem. □

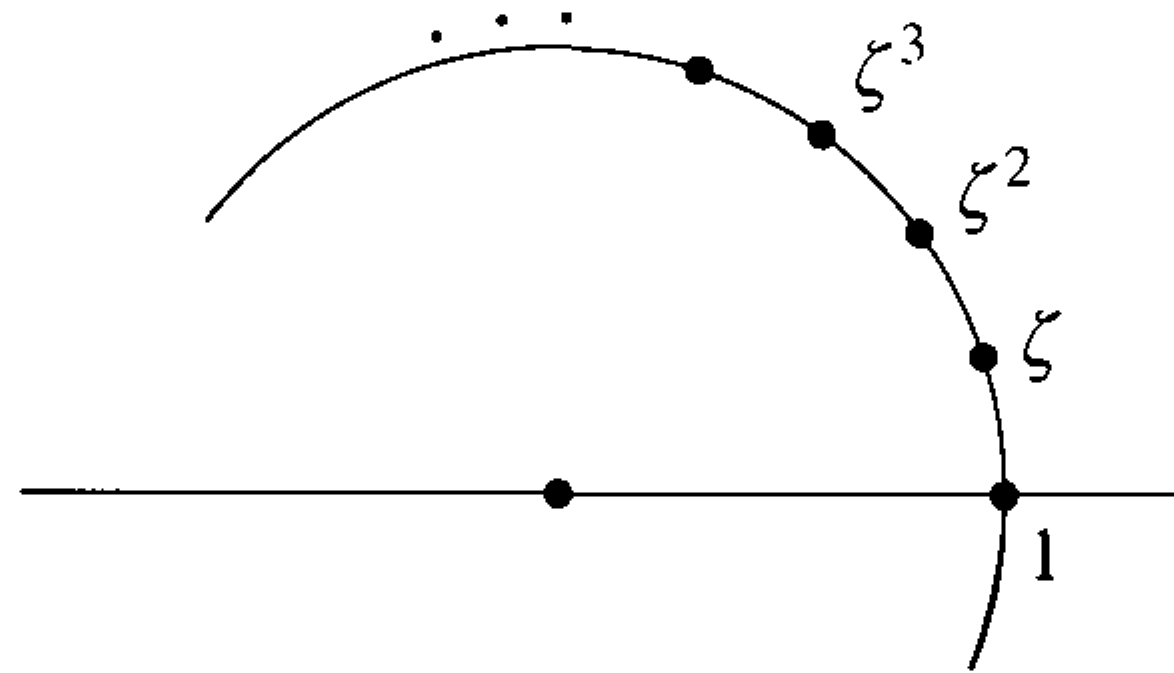
Proof of Theorem (6.4c). By an *n-th root of unity* in a field F, we mean an element α whose nth power is 1. Thus α is an nth root of unity if and only if it is a root of the polynomial

$$x^n - 1, \tag{6.16}$$

or if and only if its order, as an element of the multiplicative group $F^\times$, divides n. The nonzero elements of a finite field with q elements are $(q - 1)$-st roots of unity.

In the field of complex numbers, the nth roots of unity form a cyclic group of order n, generated by

(6.17) $$\zeta_n = e^{2\pi i/n}:$$



A field need not have many roots of unity. For example, the only real ones are ± 1. But one property of the complex numbers carries over to arbitrary fields: The nth roots of unity in any field form a cyclic group. For example, in the field $K = \mathbb{F}_4$ of order 4, the group $K^\times$ is a cyclic group of order 3, generated by α. [See (6.3).]

(6.18) **Proposition.** Let F be a field, and let H be a finite subgroup of the multiplicative group $F^\times$, of order n. Then H is a cyclic group, and it consists of all the nth roots of unity in F.

Proof. If H has order n, then the order of an element α of H divides n, so α is an nth root of unity, a root of the polynomial $x^n - 1$. This polynomial has at most n roots, so there aren't any other roots in F [Chapter 11 (1.18)]. It follows that H is the set of all nth roots of unity in F.

It is harder to show that H is cyclic. To do so, we use the Structure Theorem for abelian groups, which tells us that H is isomorphic to a direct product of cyclic groups:

$$H \approx \mathbb{Z}/(d_1) \times \cdots \times \mathbb{Z}/(d_k),$$

where $d_1 | d_2 \cdots | d_k$ and $n = d_1 \cdots d_k$. The order of any element of this product divides d_k because d_k is a common multiple of all the integers d_i. So every element of H is a root of

$$x^{d_k} - 1.$$

This polynomial has at most d_k roots in F. But H contains n elements, and $n = d_1 \cdots d_k$. The only possibility is that $n = d_k$, $k = 1$, and H is cyclic. □

Proof of Theorem (6.4a). We need to prove the existence of a field with q elements. Since we have already proved part (d) of the theorem, we know that the elements of a field of order q are roots of the polynomial $x^q - x$. Also, there exists a field L containing $\mathbb{F}_p$ in which this polynomial (or any given polynomial) factors into linear factors (5.3). The natural thing to try is to take such a field L and hope for the best—that the roots of $x^q - x$ form the subfield K of L we are looking for. This is shown by the following proposition:

(6.19) **Proposition.** Let p be a prime, and let $q = p^r$.

(a) The polynomial $x^q - x$ has no multiple root in any field L of characteristic p.
(b) Let L be a field of characteristic p, and let K be the set of roots of $x^q - x$ in L. Then K is a subfield.

This proposition, combined with Proposition (5.3), proves the existence of a field with q elements.

Proof of Proposition (6.19). (a) The derivative of $x^q - x$ is $qx^{q-1} - 1$. In characteristic p, the coefficient q is equal to 0, so the derivative is equal to -1. Since the constant polynomial -1 has no root, $x^q - x$ and its derivative have no common root! Proposition (5.7) shows that $x^q - x$ has no multiple root.

(b) Let $\alpha, \beta \in L$ be roots of the polynomial $x^q - x$. We have to show that $\alpha \pm \beta$, $\alpha\beta$, and α^{-1} (if $\alpha \neq 0$) are roots of the same polynomial. This is clear for the product and quotient: If $\alpha^q = \alpha$ and $\beta^q = \beta$, then $(\alpha\beta)^q = \alpha\beta$ and $(\alpha^{-1})^q = \alpha^{-1}$. It is not obvious for the sum, and to prove it we use the following proposition:

(6.20) **Proposition.** Let L be a field of characteristic p, and let $q = p^r$. Then in the polynomial ring $L[x,y]$, we have $(x + y)^q = x^q + y^q$.

Proof. We first prove the proposition for the case $q = p$. We expand $(x + y)^p$ in $\mathbb{Z}[x,y]$, obtaining

$$(x + y)^p = x^p + \binom{p}{1}x^{p-1}y + \binom{p}{2}x^{p-2}y^2 + \cdots + \binom{p}{p-1}xy^{p-1} + y^p,$$

by the Binomial Theorem. The binomial coefficient $\binom{p}{r}$ is an integer, and if $0 < r < p$, it is divisible by p [see the proof of (4.6) in Chapter 11]. It follows that the map $\mathbb{Z}[x,y] \longrightarrow L[x,y]$ sends these coefficients to zero and that $(x + y)^p = x^p + y^p$ in $L[x,y]$.

We now treat the general case $q = p^r$ by induction on r: Suppose that the proposition has been proved for integers less than r and that $r > 1$. Let $q' = p^{r-1}$. Then by induction, $(x + y)^q = ((x + y)^{q'})^p = (x^{q'} + y^{q'})^p = (x^{q'})^p + (y^{q'})^p = x^q + y^q$. □

To complete the proof of Proposition (6.19), we evaluate x, y at α, β to conclude that $(\alpha + \beta)^q = \alpha^q + \beta^q$. Then if $\alpha^q = \alpha$ and $\beta^q = \beta$, we find $(\alpha + \beta)^q = \alpha + \beta$, as required. The case of $\alpha - \beta$ follows by substituting $-\beta$ for β. □

Proof of Theorem (6.4b). Let K and K' be fields of order q, and let α be a generator of the cyclic group $K^\times$. Then K is certainly generated as a field extension of $F = \mathbb{F}_p$ by the element α: $K = F(\alpha)$. Let $f(x)$ be the irreducible polynomial for α over F, so that $K \approx F[x]/(f)$ (2.6). Then α is a root of two polynomials: $f(x)$ and $x^q - x$. Since f is irreducible, it divides $x^q - x$ (5.4e). We now go over to the second field K'. Since $x^q - x$ factors into linear factors in K', f has a root α' in K'.

Then $K \approx F[x]/(f) \approx F(\alpha')$. Since K and K' have the same order, $F(\alpha') = K'$; hence K and K' are isomorphic. □

Proof of Theorem (6.4e). Let $f(x)$ be an irreducible polynomial of degree r in $F[x]$, where $F = \mathbb{F}_p$ as before. It has a root α in some field extension L of F, and the subfield $K = F(\alpha)$ of L has degree r over F (3.2). Therefore K has order $q = p^r$, and by part (d) of the theorem, α is also a root of $x^q - x$. Since f is irreducible, it divides $x^q - x$, as required.

In order to prove the same thing for irreducible polynomials whose degree k divides r, it suffices to prove the following lemma:

(6.21) **Lemma.** Let k be an integer dividing r, say $r = ks$, and let $q = p^r$, $q' = p^k$. Then $x^{q'} - x$ divides $x^q - x$.

For if f is irreducible of degree k, then, as above, f divides $x^{q'} - x$, which in turn divides $x^q - x$ in $F[x]$, for any field F.

Proof of the lemma. This is tricky, because we will use the identity

$$y^d - 1 = (y - 1)(y^{d-1} + \cdots + y + 1) \tag{6.22}$$

twice. Substituting $y = q'$ and $d = s$ shows that $q' - 1$ divides $q - 1 = q'^s - 1$. Knowing this, we can conclude that $x^{q'-1} - 1$ divides $x^{q-1} - 1$ by substituting $y = x^{q'-1}$ and $d = (q - 1)/(q' - 1)$. Therefore $x^{q'} - x$ divides $x^q - x$ too. □

So we have shown that every irreducible polynomial whose degree divides r is a factor of $x^q - x$. On the other hand, if f is irreducible and if its degree k doesn't divide r, then since $[K : F] = r$, f doesn't have a root in K, and therefore f doesn't divide $x^q - x$. □

Proof of Theorem (6.4f). If k does not divide r, then $q = p^r$ is not a power of $q' = p^k$, so a field of order q can not be an extension of a field of order q'. On the other hand, if k does divide r, then Lemma (6.21) and part (d) of the theorem show that the polynomial $x^{q'} - x$ has all its roots in a field K of order q. Now Proposition (6.19) shows that K contains a field with q' elements. □

This completes the proof of theorem 6.4.

7. FUNCTION FIELDS

In this section we take a look at *function fields*, the third class of field extensions mentioned in Section 1. The field $\mathbb{C}(x)$ of rational functions in one variable x will be denoted by F throughout the section. Its elements are fractions $g(x) = p(x)/q(x)$ of polynomials $p, q \in \mathbb{C}[x]$, with $q \neq 0$. We usually cancel common factors in p and q so that they have no root in common.

Let us use the symbol P to denote the complex plane, with the complex coordinate x. A rational function $g = p/q$ determines a complex-valued function of x,

which is defined for all $x \in P$ such that $q(x) \neq 0$, that is, except at the roots of the polynomial q. Near a root of q, the function defined by g tends to infinity. These roots are called *poles* of g. (We usually use the phrase "rational function" to mean an element of the field of fractions of the polynomial ring. It is unfortunate that the word *function* is already there. This prevents us from modifying the phrase in a natural way when referring to the actual function defined by such a fraction. The terminology is ambiguous, but this can't be helped.)

A minor complication arises because formal rational functions do not define functions at certain points, namely at their poles. When working with the whole field F, we have to face the fact that every value α of x can be a pole of a rational function, for example of the function $(x - \alpha)^{-1}$. There is no way to choose a common domain of definition for all rational functions at once. Fortunately this is not a serious problem, and there are two ways to get around it. One is to introduce an extra value ∞ and to define $g(\alpha) = \infty$ if α is a pole of g. This is actually the better way for many purposes, but for us another way will be easier. It is simply to ignore bad behavior at a finite set of points.

Any particular computations we may make will involve finitely many functions, so they will be valid except at a finite set of points of the plane P, the poles of these functions. A rational function is determined by its value at any infinite set of points. This is proved below, in Lemma (7.2). So we can throw finite sets out of the domain of definition as needed, without losing control of the function. Since a rational function is continuous wherever it is defined, we can recover its value at a point x_0 which was thrown out unnecessarily, as

(7.1) $$g(x_0) = \lim_{x \to x_0} g(x).$$

(7.2) **Lemma.** If two rational functions f_1, f_2 agree at infinitely many points of the plane, then they are equal elements of F.

Proof. Say that $f_i = p_i/q_i$, where $p_i, q_i \in \mathbb{C}[t]$. Let $h(x) = p_1q_2 - p_2q_1$. If $h(x)$ is the zero polynomial, then $f_1 = f_2$. If $h(x)$ is not zero, then it has finitely many roots, so there are only finitely many points at which $f_1 = f_2$. □

In order to formalize the intuitive procedure of ignoring trouble at finite sets of points, it is convenient to have a notation for the result of throwing out a finite set. Given an infinite set U, we will denote by U' a set obtained from U by deleting an unspecified finite subset, which is allowed to vary as needed:

(7.3) $$U' = U - (\text{variable finite set}).$$

By a *function* on U' we mean an equivalence class of complex-valued functions, each defined except on a finite subset of U. Two such functions f, g are called *equal on* U' if there is a finite subset Δ of U such that f and g are defined and equal on $U - \Delta$. (We could also refer to this property by saying that $f = g$ *almost everywhere* on U. However, in other contexts, "almost everywhere" often means "except

on a set of measure zero," rather than "except on a finite set.") A function f on U' will be called *continuous* if it is represented by a continuous function on some set $U - \Delta$.

The set of continuous functions on U' will be denoted by

$$\mathcal{F}(U) = \{\text{continuous functions on } U'\}. \tag{7.4}$$

This set forms a ring, with the usual laws of addition and multiplication of functions:

$$[f + g](x) = f(x) + g(x) \quad \text{and} \quad [fg](x) = f(x)g(x). \tag{7.5}$$

Lemma (7.2) has the following corollary:

(7.6) **Proposition.** The field $F = \mathbb{C}(x)$ is isomorphic to a subring of the ring $\mathcal{F}(P)$, where P is the complex plane. □

Let us now examine one of the simplest function fields in more detail. We are going to need polynomials with coefficients in the field F. Since the symbol x has already been assigned, we use y to denote the new variable. We will study the quadratic field extension K obtained from F by adjoining a root of $f(y)$, where $f = y^2 - x$. Since f depends on the variable x as well as on y, we will also write

$$f = f(x,y) = y^2 - x. \tag{7.7}$$

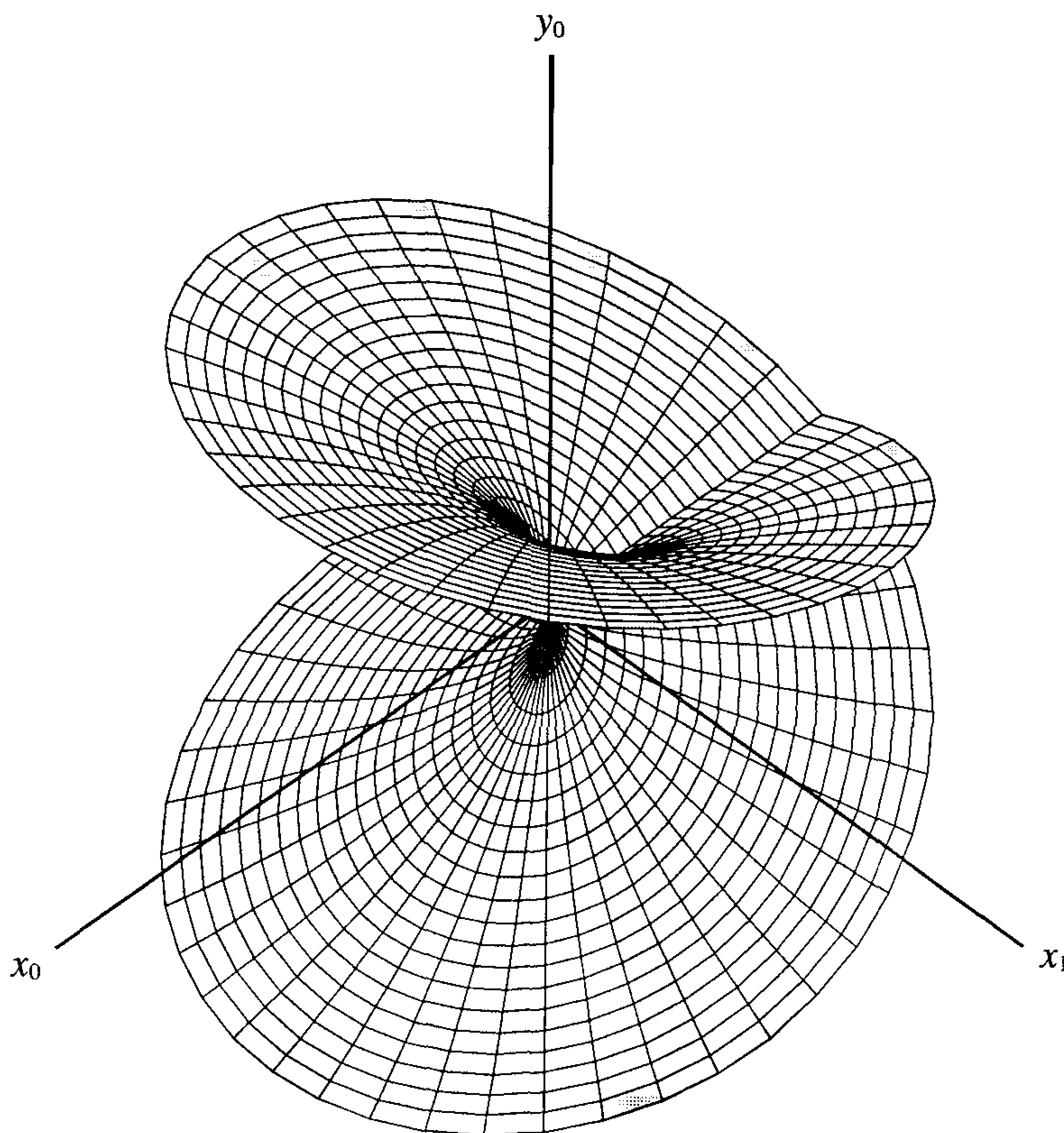
The polynomial $y^2 - x$ is an irreducible element of $F[y]$, so K can be constructed as the abstract field $F[y]/(f)$. The residue of the variable y is a root of f in K.

The importance of function fields comes from the fact that their elements can be interpreted as actual functions. In our case, we can define a square root *function* h, by choosing one of the two values of the square root for each complex number $x : h(x) = \sqrt{x}$. Then h can be interpreted as a function on P'. However, since there are two values of the square root whenever $x \neq 0$, we need to make a lot of choices to define this function. This isn't very satisfactory. If x is real and positive, it is natural to choose the positive square root, but no choice will give a continuous function on the whole complex plane.

The locus S of solutions of the equation $y^2 - x = 0$ in $\mathbb{C}^2$ is called the *Riemann surface* of the polynomial $y^2 - x$ (see Section 8 of Chapter 10). It is depicted below in Figure (7.9), but in order to obtain a surface in real 3-space, we have dropped one coordinate. The complex two-dimensional space $\mathbb{C}^2$ is identified with $\mathbb{R}^4$ by the usual rule $(x,y) = (x_0 + x_1 i, y_0 + y_1 i) \longleftrightarrow (x_0, x_1, y_0, y_1)$. The figure depicts the locus

$$\{(x_0, x_1, y_0) \mid y_0 = \text{real part of } (x_0 + x_1 i)^{1/2}\}. \tag{7.8}$$

This is a projection of S from $\mathbb{R}^4$ to $\mathbb{R}^3$.



(7.9) **Figure.** The Riemann surface $y^2 = x$.

The Riemann surface S does not cut itself along the negative x_0-axis as the projected surface does. Every negative real number x has two purely imaginary square roots, but the real parts of these square roots are zero. This produces the apparent self-crossing in the projected surface. Actually, S is a two-sheeted branched covering of P, as defined in Chapter 10 (8.13), and the only branch point is at $x = 0$.

Figure (7.9) shows the problem encountered when we try to define the square root as a single-valued function. When x is real and positive, the positive square root is the natural choice. We would like to extend this choice continuously over the complex plane, but we run into trouble: Winding once around the origin in complex x-space brings us back to the negative square root. It is better to accept the fact that the square root, as a solution of the equation $y^2 - x = 0$, is a multi-valued function on P'.

Now there is an amazing trick which will allow us to solve any polynomial equation $f(x,y) = 0$ with a *single-valued* function, without making arbitrary choices. The trick is to replace the complex plane P by the Riemann surface S, the locus $f(x,y) = 0$. We are given two functions on S, namely the restrictions of the

coordinate functions on $\mathbb{C}^2$. In order to keep things straight, let us introduce new symbols for these functions, say X, Y:

(7.10) $$\mathrm{X}(x,y) = x \quad \text{and} \quad \mathrm{Y}(x,y) = y, \quad \text{for } (x,y) \in S.$$

These restrictions of the coordinate functions to S are related by the equation $f(\mathrm{X},\mathrm{Y}) = 0$, because by definition of S, $f(x,y) = 0$ at any point of S.

(7.11) **Proposition.** Let $f(x,y)$ be an irreducible polynomial in $\mathbb{C}[x,y]$ which is not a polynomial in x alone, and let $S = \{(x,y) \mid f(x,y) = 0\}$ be its Riemann surface. Let $K = F[y]/(f)$ be the field extension defined by f. Then K is isomorphic to a subring of the ring $\mathcal{F}(S)$ of continuous functions on S'.

Proof. Let $g(x)$ be a rational function. Since X is the restriction of a coordinate function on $\mathbb{C}^2$, the composed function $g(\mathrm{X})$ is continuous on S except at the points which lie above the poles of g. There are finitely many such points [Chapter 10 (8.11)]. So $g(\mathrm{X})$ is a continuous function on S'. We define a homomorphism $F \longrightarrow \mathcal{F}(S)$ by sending $g(x)$ to $g(\mathrm{X})$. Next, the Substitution Principle extends this map to a homomorphism

(7.12) $$\varphi: F[y] \longrightarrow \mathcal{F}(S),$$

by sending $y \rightsquigarrow \mathrm{Y}$. Since $f(\mathrm{X},\mathrm{Y}) = 0$, the polynomial $f(x,y)$ is in the kernel of φ. Since $K = F[y]/(f)$, the mapping property of quotients [Chapter 10 (4.2)] gives us a map $\overline{\varphi}: K \longrightarrow \mathcal{F}(S)$ which sends the residue of y to Y. Since K is a field, $\overline{\varphi}$ is injective. □

(7.13) **Definition.** An *isomorphism* of branched coverings S_1, S_2 of the plane P is a homeomorphism $\varphi': S_1' \longrightarrow S_2'$ which is compatible with the maps $\pi_i: S_i \longrightarrow P$, that is, such that $\pi_2'\varphi = \pi_1'$:

$$\begin{array}{ccc} S_1' & \xrightarrow{\varphi'} & S_2'. \\ & \searrow^{\pi_1'} \quad {}^{\pi_2'}\swarrow & \\ & P & \end{array}$$

By this we mean that φ' is defined except on a finite set of S_1 and that when suitable finite sets are omitted from S_1 and S_2, φ' is a homeomorphism.

A branched covering S is called *connected* if the complement S' of an arbitrary finite set of S is a path-connected set.

We will now state a beautiful theorem which describes the finite extensions of the field of rational functions. Let $\mathcal{E}_n$ denote the set of isomorphism classes of extension fields K of F of degree n. Let $\mathcal{C}_n$ denote the set of isomorphism classes of connected n-sheeted branched coverings $\pi: S \longrightarrow P$ of the plane.

(7.14) **Theorem.** *Riemann Existence Theorem:* There is a bijective map $\Phi_n: \mathcal{E}_n \longrightarrow \mathcal{C}_n$. If K is the extension obtained by adjoining a root of an irreducible

polynomial $f(x, y) \in \mathbb{C}[x, y]$, then the class of branched coverings corresponding to K is represented by the Riemann surface of f. □

The proof of this theorem is a suitable topic for a course in complex variables. It requires too much analysis to give here. Using it, however, we can associate a branched covering of the plane, unique up to isomorphism, to every finite extension field K of F. This covering is called the *Riemann surface of the extension field K.* The Riemann surface of F is the complex plane P itself.

Here are two striking corollaries of the theorem:

(7.15) **Corollary.** Given a connected n-sheeted branched covering S of the plane, there is a polynomial $f(x, y)$ of degree n in y whose Riemann surface is isomorphic to S.

This follows from the surjectivity of the map Φ_n and from a fact which will be proved in the next chapter [Chapter 14 (4.1)], that every finite extension K of F can be obtained by adjoining a single element. □

(7.16) **Corollary.** Let f, g be irreducible polynomials in $\mathbb{C}[x, y]$, with Riemann surfaces S, T. Let α be a root of $f(y)$ in an extension field of F. If S and T are isomorphic branched coverings, then $g(y)$ has a root in $F(\alpha)$.

This follows from the injectivity of the map Φ_n. □

Visualization of Riemann surfaces is complicated by the fact that they are embedded in $\mathbb{C}^2$, a four-dimensional real space. One aid to constructing and visualizing them is a method known as *cut and paste*. If we cut the surface $y^2 - x$ open along the negative real axis, the double locus in Figure (7.9), then it decomposes into the two parts $\operatorname{re} Y > 0$ and $\operatorname{re} Y < 0$. Each of these parts projects to the x-plane P in a bijective way, if we disregard what happens along the cut. Turning this procedure around, we can construct a surface which is homeomorphic to S in the following way: We stack two copies P_1, P_2 of the complex plane over P and cut them open along the negative real axis $(-\infty, 0]$. These copies of P are called *sheets*. Then we glue side A of P_1 to side B of P_2 and vice versa (see below). Four dimensions are needed to embed S without crossings.

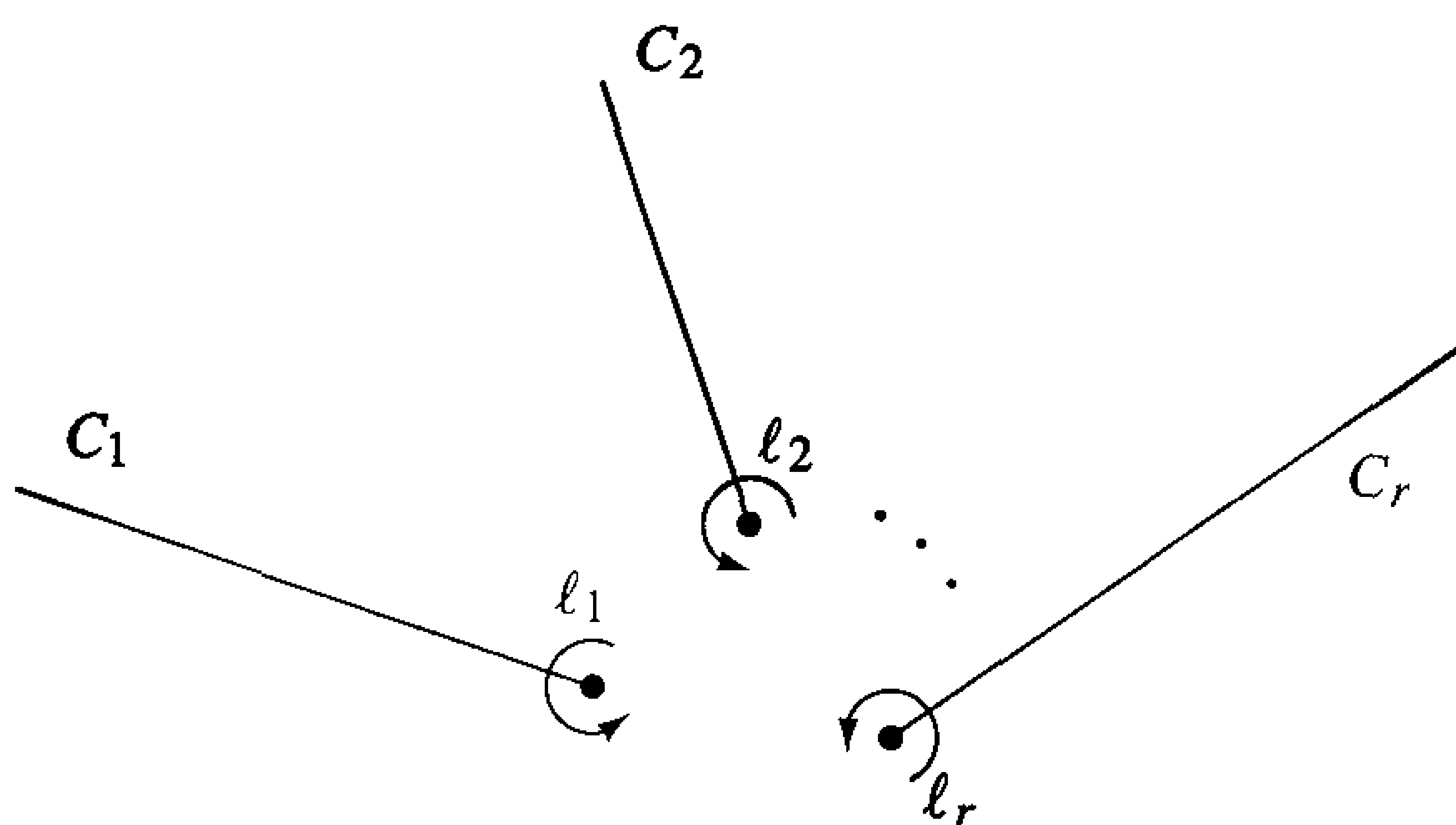
side A of cut
______________________.
side B of cut

(7.17) **Figure.**

To construct a general branched covering S of the plane by the cut-and-paste procedure, we begin with n copies of the plane P, called *sheets*. The sheets are labelled $P_1, \ldots, P_n$ and are stacked up over P. We also select a finite set of points $\alpha_1, \ldots, \alpha_r$ of P to be branch points. For each branch point α_ν, we choose a curve C_ν

beginning at α_ν and going to infinity in an arbitrary direction. This should be done in such a way that the curves C_ν do not intersect. The sheets P_i are cut open along these curves. Then various sheets are glued to others along opposite edges of the cuts.

To describe the resulting covering S, we need only describe the permutations σ_ν by which the sheets are glued together along the cuts. To be specific, we draw a small loop ℓ_ν around the point α_ν in the counterclockwise direction. Then if the permutation σ_ν sends the index **1** to **3**, we glue sheet P_1 to sheet P_3 as we cross C_ν. This means that if we start on sheet P_1 and wind once around the loop ℓ_ν, we return on sheet P_3. The permutation σ_ν can be arbitrary.



The points α_ν are called *branch points* of the surface S because the surface decomposes into n disjoint sheets near any other point of P. It won't have n disjoint sheets above the point α_ν unless the permutation σ_ν is the identity. If $\sigma_\nu = 1$, then each sheet is glued back to itself along the cut C_ν, so that cut was not needed. But it is convenient to allow this as a possibility. Let's call α_ν a *true* branch point if $\sigma_\nu \neq 1$. Some of the points α_ν may not be true branch points. However, all true branch points are among them.

It is important to note that the numbering of the sheets is arbitrary and, in particular, that the concept of a "top sheet" has no intrinsic meaning for the Riemann surface of a polynomial. If there was a top sheet, we could define y as a single-valued function by choosing the value on that sheet. One can do this only once the Riemann surface has been cut open. This is the whole point; wandering around on the surface will lead us from one sheet to another.

It is not difficult to decide when two such branched coverings are isomorphic.

(7.18) **Proposition.** Let S, T be branched coverings which are constructed as above, with the same branch points α_ν and the same curves C_ν, but using different sets of permutations $(\sigma_1, \ldots, \sigma_r)$ and $(\tau_1, \ldots, \tau_r)$. Then S and T are isomorphic coverings if and only if the two sets of permutations are conjugate, that is, if and only if there is a permutation ρ such that $\tau_\nu = \rho^{-1}\sigma_\nu\rho$ for all ν.

Proof. Let σ, C stand for σ_ν, C_ν. Our rule is that P_i is glued to $P_{i\sigma}$ along C. Suppose that we relabel the sheets $P_1, \ldots, P_n$, changing the numbers by a permutation ρ. To keep old and new labellings straight, let's label the renumbered sheets as Q_j. So for every i, P_i is relabelled as $Q_{i\rho}$. The rule now tells us to glue $P_i = Q_{i\rho}$ to $P_{i\sigma} = Q_{i\sigma\rho}$. Substituting $i = j\rho^{-1}$ shows that the rule glues Q_j to $Q_{j\rho^{-1}\sigma\rho}$. Thus the permutation which describes this gluing rule is the conjugate $\rho^{-1}\sigma_\nu\rho$ of the old permutation σ_ν. Since the covering is not changed by the relabelling process, this shows that a conjugate set of permutations defines an isomorphic covering.

Conversely, let $\varphi: S \longrightarrow T$ be an isomorphism of coverings. Let $P_1, \ldots, P_n$ be the sheets which are used to construct S, and let $Q_1, \ldots, Q_n$ be those used to construct T. Then since P_i is connected and since T, when cut open, is a disjoint union of the open sets Q_j, the image of P_i must be contained in a single sheet Q_j. Since φ is compatible with the projections to P, which are homeomorphisms except on the cuts, the restriction of φ to P_i must be a bijection onto the sheet Q_j. So we can renumber the sheets Q_j so that P_i is mapped to Q_i. This changes the permutations τ_ν to conjugates, as above. So we may assume that φ carries P_i to Q_i. Also, φ is continuous across the cuts. Therefore if crossing the cut C_ν on sheet P_i leads to P_j, then, similarly, crossing on Q_i must lead to Q_j. Therefore $\sigma_\nu = \tau_\nu$. □

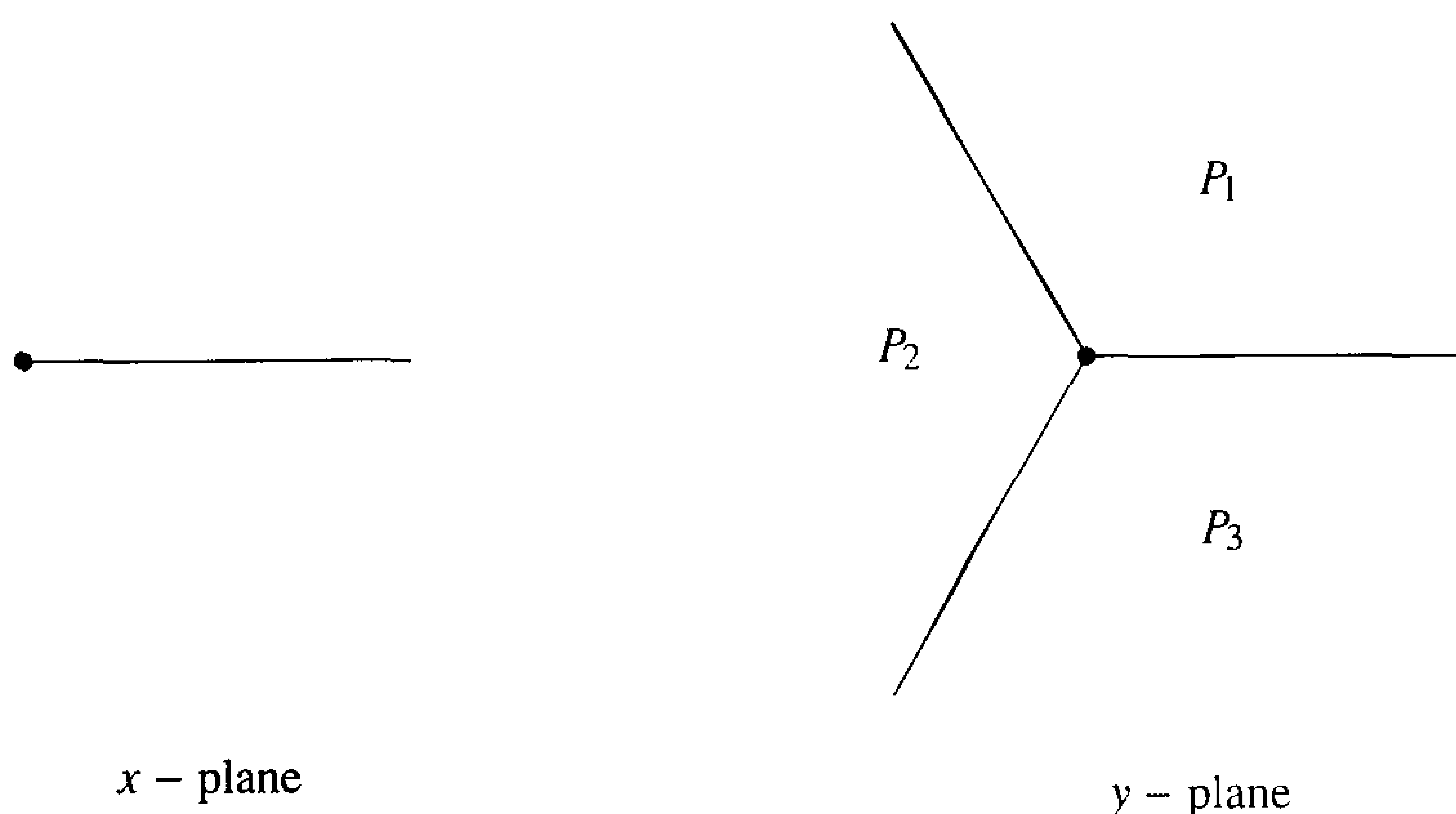
We can also start with an arbitrary branched covering S and reconstruct it in this way: Say that S is branched at the points $\alpha_1, \ldots, \alpha_r \in P$. As above, we choose nonintersecting curves C_i beginning at α_i and going to infinity. Then if S is cut open above the curves C_i, it decomposes into n sheets. This is a theorem of topology, because the complement of the curves C_i in P is simply connected [Munkres, *Topology* p. 342, exc. 8]. Therefore a covering homeomorphic to S can be reconstructed from n sheets $P_1, \ldots, P_n$ by cutting them open along the curves and gluing together to mix up the sheets.

We will now describe the Riemann surfaces of a few simple polynomials f. This is usually difficult to do when f is complicated.

(7.19) **Example.** The Riemann surface of $y^3 - x$: Here y represents a cube root of x, and S is a three-sheeted covering of P. The only branch point is $x = 0$. We cut S open above the positive real axis $C = [0, \infty]$. This decomposes S into three sheets P_1, P_2, P_3, and it is reasonable to guess that the gluing along the cut is done by a cyclic permutation.

This case is fairly easy to analyze because x is a single-valued function of y. Because of this, we can interpret S as the graph of a function from y-space to x-space, which implies that the projection of S onto the complex y-plane is bijective. We identify S with the y-plane using this projection and cut it open above C. This will decompose the plane into three parts corresponding to the sheets P_i. The rules for gluing will be evident when this decomposition is made explicit.

The values of y lying over the cut C are those for which $y^3 = x$ is real and positive. They are $y = re^{i\theta}$, where $\theta = 0$, $2\pi/3$, or $4\pi/3$. So the sheets are sectors.

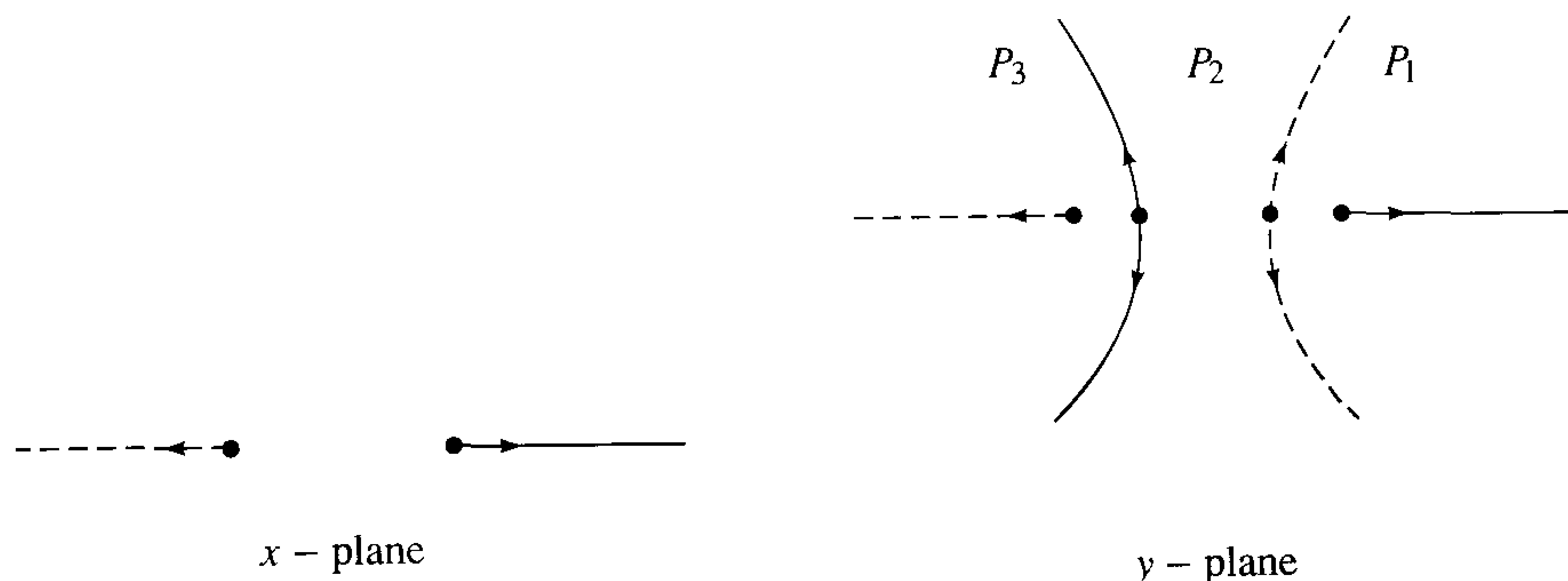


In the figure, the sectors have been numbered arbitrarily. Note that under the map $y \rightsquigarrow y^3 = x$, each of the three sectors is stretched radially and maps bijectively to the entire plane, disregarding the cuts. As we move along S to cross the cut in the x-plane, we also cross one of the three cuts in the y-plane. As predicted, this permutes the sheets by the cyclic permutation $(\mathbf{1\,2\,3})$. □

(7.20) **Example.**

The Riemann surface of $f(x, y) = y^3 - 3y - x$: The points x at which this polynomial has fewer than three roots are found by solving the equations $f = \partial f/\partial y = 0$ [see Chapter 10 (8.12)]. Here $\partial f/\partial y = 3(y^2 - 1)$. So the solutions are $y = \pm 1$, and hence $x = \pm 2$. We may cut S open above the curves $C_1 = (-\infty, -2]$ and $C_2 = [2, \infty)$, to decompose it into three sheets.

Again, x is a single-valued function of y, and we can analyze the gluing of the sheets by cutting the y-plane apart suitably. To do so, we ask for the values of y such that x lies on one of the curves C_i. Since these curves are on the real x-axis, we begin by solving the equation $\operatorname{im} x = 0$. Setting $y = u + vi$, we find $\operatorname{im} x = \operatorname{im}(y^3 - 3y) = v(3u^2 - v^2 - 3)$. The solutions are the u-axis $v = 0$ and the two branches of the hyperbola $3u^2 - v^2 = 3$. The points on the u-axis in the interval $(-2, 2)$ correspond to $x \in (-2, 2)$, so they do not lie over the cuts.

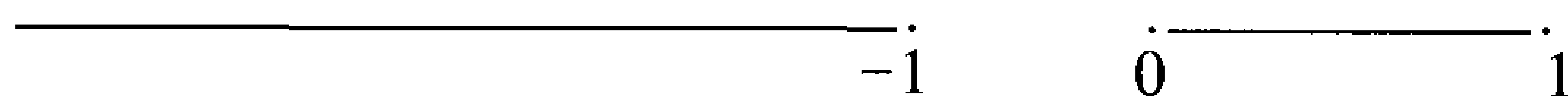


Again, each of the three regions into which the y-plane decomposes is mapped bijectively to the x-plane by the function $y^3 - 3y$, disregarding the cut as always. In the figure, the dotted curves are those which lie over C_1. The figure shows that moving on S to cross the curve $(-\infty, -2]$ interchanges the sheets P_1, P_2, leaving P_3 alone, and similarly that crossing above $[2, \infty)$ interchanges P_2, P_3. So the branching is described by the transposition $(\mathbf{2\,3})$ at the branch point $x = -2$ and by $(\mathbf{1\,2})$ at $x = 2$. □

(7.21) **Example.** The Riemann surface of $y^2 - x^3 + x^2$: There are two points $x = 0, 1$ above which S has fewer than two points. However, at $x = 0$ the sheets cross without getting mixed up, so the only true branch point is $x = 1$. To see this we make the change of variable $x = x$, $z = y/x$, which is defined and invertible except at $x = 0$. Then $z^2 - x + 1 = 0$. The given surface S becomes homeomorphic to the Riemann surface of $z^2 - x + 1$ when the points above the origin are deleted, and the surface can be reduced to (7.9) by a translation in the x-plane. □

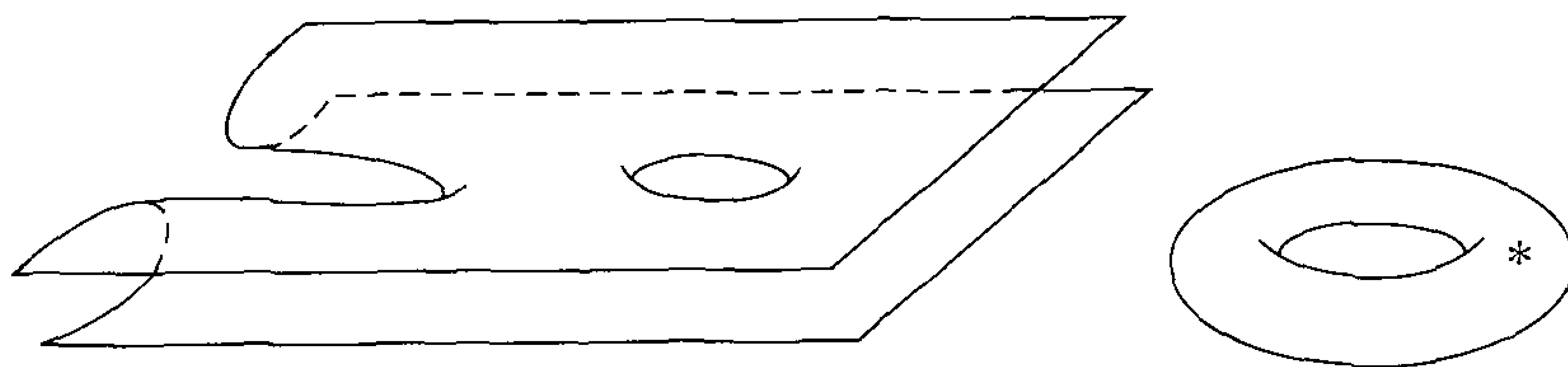
When it is not possible to solve for x as a single-valued function of y, the problem of describing the gluing data becomes more difficult. We will work out one example of this type.

(7.22) **Example.** The Riemann surface of $y^2 - (x^3 - x)$: There are three points at which $x^3 - x = 0$, namely $x = 0, \pm 1$, and the surface has three branch points at which it behaves like the Riemann surface of $y^2 - x$ at the origin. Our systematic procedure is to make cuts from these three branch points to infinity, but in this case another choice of cuts is easier to analyze. The values of x such that y is purely imaginary are the real x such that $x^3 - x \leq 0$. These are the points in the two intervals $(-\infty, -1]$ and $[0, 1]$. If we cut S open along these two intervals, it will decompose into the parts $\operatorname{re} y > 0$ and $\operatorname{re} y < 0$. Thus we can reconstruct the surface S by stacking up two copies of P, cutting them open along the intervals and gluing to mix up the sheets as before.



(7.23) **Figure.**

The fact that a surface constructed by the cut-and-paste method crosses itself along the cuts makes it confusing to visualize directly. But since the cuts are along the real axis in this example, we can avoid crossings by turning one of the sheets over. This ruins the representation of S as a double covering of P, but the advantage is that the sheets are now glued along the same side of the cut. There are two such cuts in Figure (7.23). Turning one sheet over and stretching to pull the slits apart after gluing results in the following picture: This Riemann surface is homeomorphic to a torus with one point deleted. □



8. TRANSCENDENTAL EXTENSIONS

In this section we will take a brief look at some transcendental field extensions. We saw in Propositon (2.5) that the structure of the field extension $F(\alpha)$ generated by a single transcendental element α over a field F does not depend on the element α. But if two transcendental elements α, β are adjoined at the same time, the structure of the field $F(\alpha, \beta)$ which is obtained will depend on whether or not the elements α and β are algebraically related, and if they are related, the structure will depend on the nature of this relation. For example, $\alpha = \sqrt{\pi}$ and $\beta = \sqrt[4]{\pi}\sqrt{\pi - 1}$ are transcendental numbers over $\mathbb{Q}$, which are related by the equation

$$\beta^2 - \alpha^3 + \alpha = 0.$$

In general, we call a set of elements $\{\alpha_1, \ldots, \alpha_n\}$ of an extension field $K \supset F$ *algebraically dependent over* F if there is a nonzero polynomial in n variables $f(x_1, \ldots, x_n) \in F[x_1, \ldots, x_n]$ such that

$$f(\alpha_1, \ldots, \alpha_n) = 0,$$

and we call them *algebraically independent over* F if there is no such polynomial. Thus $\sqrt{\pi}$ and $\sqrt[4]{\pi}\sqrt{\pi - 1}$ are algebraically dependent over $\mathbb{Q}$. It is conjectured that e and π are algebraically independent, but this has not been proved.

We can interpret algebraic independence in terms of the substitution homomorphism $\varphi: F[x_1, \ldots, x_n] \longrightarrow K$ sending $f(x_1, \ldots, x_n) \rightsquigarrow f(\alpha_1, \ldots, \alpha_n)$. The elements $\alpha_1, \ldots, \alpha_n$ are algebraically independent if $\ker \varphi = 0$, that is, if φ is injective, and algebraically dependent otherwise. Passing to fields of fractions gives this proposition:

(8.1) **Proposition.** If $\alpha_1, \ldots, \alpha_n$ are algebraically independent, then $F(\alpha_1, \ldots, \alpha_n)$ is isomorphic to the field $F(x_1, \ldots, x_n)$ of rational functions in $x_1, \ldots, x_n$, the field of fractions of $F[x_1, \ldots, x_n]$. □

An extension of the form $F(\alpha_1, \ldots, \alpha_n)$, where α_i are algebraically independent, is called a *pure transcendental* extension.

(8.2) **Definition.** A *transcendence basis* for a field extension K of F is a set of elements $(\alpha_1, \ldots, \alpha_n)$ which are algebraically independent and such that K is an algebraic extension of the field $F(\alpha_1, \ldots, \alpha_n)$.

(8.3) **Theorem.** Let $(\alpha_1,\ldots,\alpha_m)$ and $(\beta_1,\ldots,\beta_n)$ be elements in an extension K of a field F. Assume that K is algebraic over $F(\beta_1,\ldots,\beta_n)$ and that $\alpha_1,\ldots,\alpha_m$ are algebraically independent over F. Then $m \leq n$, and $(\alpha_1,\ldots,\alpha_m)$ can be completed to a transcendence basis for K by adding $(n-m)$ of the elements β_i.

We leave the proof of this theorem as an exercise. □

(8.4) **Corollary.** Any two transcendence bases for an extension $F \subset K$ have the same number of elements. □

(8.5) **Definition.** The *transcendence degree* of K is the number of elements in a transcendence basis, or is infinite if no finite transcendence basis exists.

(8.6) **Examples**

(a) The fields $F(x_1,\ldots,x_n)$ of rational functions in n variables are not isomorphic extensions of F for different values of n, because $(x_1,\ldots,x_n)$ is a transcendence basis.

(b) Let α, β be as at the beginning of the section. The single element π forms a transcendence basis for $K = \mathbb{Q}(\alpha,\beta)$ over $\mathbb{Q}$. Therefore (8.3) implies that, as was asserted above, any two elements of K are algebraically dependent. The element β is another transcendence basis.

(c) Consider any two polynomials or rational functions in one variable $f, g \in F(x)$. There is a nonzero polynomial $\varphi(y,z) \in F[y,z]$ such that $\varphi(f,g) = 0$. For, the transcendence degree of $F(x)$ is 1, and hence f, g are algebraically dependent.

Most field extensions aren't pure transcendental, though this may be difficult to decide for a particular extension. Here are two examples:

(8.7) **Proposition.**

(a) The function field $L = \mathbb{C}(x)[y]/(y^2 - x^3)$ is a pure transcendental extension of $\mathbb{C}$. It is the field of rational functions in $t = y/x$.

(b) The function field $K = \mathbb{C}(x)[y]/(y^2 - x^3 + x)$ is not a pure transcendental extension of $\mathbb{C}$. That is, there is no element $t \in K$ such that $K = \mathbb{C}(t)$.

Proof. In both cases, the transcendence degree of K over $\mathbb{C}$ is 1, because x is a transcendence basis.

(a) Let $t = y/x$. Then $\mathbb{C}(t) \subset L$ because $t \in L$. Now L is generated by x and y, by definition. On the other hand, $x = t^2$ and $y = t^3$. Therefore $L = \mathbb{C}(t)$. Since K has transcendence degree 1, (8.4) shows that t is transcendental.

(b) (*Sketch*) To show that K is not a field of rational functions, we appeal to the geometry of its Riemann surface. We saw in the last section that this surface is a torus from which one point has been deleted. On the other hand, the Riemann surface of the field of rational functions $\mathbb{C}(t)$ is the complex plane itself. Now, it is a theorem

of topology that the torus and the plane are not homeomorphic and that they do not become homeomorphic when finite sets are deleted. If we admit this theorem, then the next proposition will complete the proof.

(8.8) **Proposition.** Let $K = \mathbb{C}(x)[y]/(f)$ and $L = \mathbb{C}(t)[u]/(g)$ be function fields with Riemann surfaces S, T respectively. A homomorphism $\varphi: L \longrightarrow K$ which is the identity on the subfield $\mathbb{C}$ induces a map $\varphi^*: S' \longrightarrow T$ between their Riemann surfaces, which is defined and continuous except on a finite set of points of S. If φ is an isomorphism, then φ^* becomes a homeomorphism when suitable finite sets are deleted from S and T.

Note that the map φ^* goes from the Riemann surface of K to that of L, in the opposite direction from φ.

Proof. The Riemann surface T is the locus $g(t, u) = 0$ in $\mathbb{C}^2$. According to Proposition (7.11), every element $\alpha \in K$ defines a continuous function on S', so the pair of functions $(\varphi(t), \varphi(u))$ defines a continuous map $S' \longrightarrow \mathbb{C}^2$. Since $g(t, u) = 0$ in L and since φ is a homomorphism which leaves the coefficients of g fixed, $g(\varphi(t), \varphi(u)) = 0$ too. So S' is mapped to T. This is the required map φ^*. If φ is an isomorphism, its inverse defines a map $T' \longrightarrow S$ which is an inverse function to φ^* on the complement of a finite set. □

9. ALGEBRAICALLY CLOSED FIELDS

A field F is said to be *algebraically closed* if every polynomial $f(x) \in F[x]$ of positive degree has a root in F. The fact that the field $\mathbb{C}$ of complex numbers is algebraically closed is called the Fundamental Theorem of Algebra.

(9.1) **Theorem.** *Fundamental Theorem of Algebra*: Every nonconstant polynomial with complex coefficients has a complex root.

We have used this theorem often already. A proof is at the end of the section.

If a field F is algebraically closed, then every nonconstant polynomial $f(x) \in F[x]$ has a linear factor $x - \alpha$, so the only irreducible polynomials are the linear ones. Consequently every polynomial is a product of linear factors. Also, there are no algebraic extensions of F other than F itself (whence the phrase algebraically closed). For if α is algebraic over F, then it is a root of a monic irreducible polynomial $f(x) \in F[x]$. This polynomial must have the form $x - \alpha$, so $\alpha \in F$.

It may be convenient to think of a field F which is being studied as a subfield of an algebraically closed field. For instance, we like to think of number fields as subfields of $\mathbb{C}$. Let us call an extension field K of F an *algebraic closure of* F if

(9.2) (i) K is algebraic over F, and
(ii) K is algebraically closed.

(9.3) **Corollary.** Let F be a subfield of $\mathbb{C}$. The subset $\overline{F}$ of $\mathbb{C}$ consisting of all numbers which are algebraic over F is an algebraic closure of F.

Proof. The fact that $\overline{F}$ is a field has been proved (3.10). To show that $\overline{F}$ is algebraically closed, let $f(x) \in \overline{F}[x]$ be a nonconstant polynomial. Then $f(x)$ has a root α in $\mathbb{C}$, and $\overline{F}(\alpha)$ is algebraic over $\overline{F}$. Since $\overline{F}$ is algebraic over F, α is algebraic over F by (3.11). So $\alpha \in \overline{F}$. □

It is not hard to construct an algebraic closure of a finite field $\mathbb{F}_p$, as a union of the fields $\mathbb{F}_q$, where $q = p^r$ is a power of p. To do this, we choose a sequence of integers $r_1, r_2, \ldots$ with these properties: (i) r_i divides r_{i+1}, and (ii) every integer n divides some r_i. We may take $r_i = i!$, for example. We set $q_i = p^{r_i}$ and $F_i = \mathbb{F}_{q_i}$. It follows from (i) that F_{i+1} contains a subfield isomorphic to F_i (6.4), so we can build a tower of fields $F_1 \subset F_2 \subset \cdots$. Let $\overline{F}$ be the union of this chain of fields. Then (ii) tells us that every finite field $\mathbb{F}_q$, $q = p^r$, is isomorphic to a subfield of some F_i, hence to a subfield of $\overline{F}$. This field is an algebraic closure of $\mathbb{F}_p$.

The following theorem can be proved using Zorn's Lemma.

(9.4) **Theorem.** Every field F has an algebraic closure, and if K_1, K_2 are two algebraic closures of F, there is an isomorphism $\varphi: K_1 \longrightarrow K_2$ which is the identity map on the subfield F. □

Thus the algebraic closure is essentially unique.

(9.5) **Corollary.** Let $\overline{F}$ be an algebraic closure of F, and let K be any algebraic extension of F. There is a subextension $K' \subset \overline{F}$ isomorphic to K. □

Proof of the Fundamental Theorem of Algebra. To show that $f(x_0) = 0$, it is enough to show that the absolute value $|f(x_0)|$ is zero. The existence of such a value $x_0 \in \mathbb{C}$ is proved by the following two lemmas:

(9.6) **Lemma.** Let $f(x)$ be a nonconstant polynomial, and let $x_0 \in \mathbb{C}$ be a point at which $f(x_0) \neq 0$. Then $|f(x_0)|$ is not the minimum value of $|f(x)|$.

(9.7) **Lemma.** Let $f(x)$ be a complex polynomial. Then $|f(x)|$ takes on a minimum value at some point $x_0 \in \mathbb{C}$.

Proof of Lemma (9.6). We first note that the polynomial $x^k - c$ has a root for all $c \in \mathbb{C}$. A nonnegative real number r has a real kth root because the continuous function x^k, which is zero when $x = 0$ and large when x is a large real number, takes on all real values ≥ 0, by the Intermediate Value Theorem. We write the complex number c in the form $c = re^{i\theta}$, where $r = |c|$ and $\theta = \arg c$. Let s be a real kth root of r. Then the required kth root of c is

$$\alpha = se^{i\theta/k}. \tag{9.8}$$

Now let $f(x)$ be a nonconstant polynomial, and let $x_0 \in \mathbb{C}$ be a point at which $f(x_0) \neq 0$. It is convenient to normalize f. We make a change of variable, replacing

x by $x + x_0$, to shift the point in question to the origin: $x_0 = 0$. We also multiply $f(x)$ by $f(0)^{-1}$. Then $f(0) = 1$, and we must show that 1 is not the minimum value of $|f(x)|$.

Let k denote the lowest nonzero power of x occurring in f, so that

$$f(x) = 1 + ax^k + (\text{terms of degree} > k).$$

Let α be a kth root of $-a^{-1}$. We make a final change of variable, replacing x by αx. Then f takes the form

$$f(x) = 1 - x^k + (\text{higher-degree terms}) = 1 - x^k + x^{k+1}g(x),$$

for some polynomial $g(x)$. For small positive real x, the Triangle Inequality shows that

$$|f(x)| \le |1 - x^k| + |x^{k+1}g(x)| = 1 - x^k + x^{k+1}|g(x)| = 1 - x^k(1 - x|g(x)|).$$

Since $x|g(x)|$ is small for small x, the term $x^k(1 - x|g(x)|)$ is positive when x is a sufficiently small positive real number. For such x, $|f(x)| < |f(0)|$. □

Proof of Lemma (9.7). We may assume that f is not a constant polynomial. For large x, $f(x)$ is also large:

(9.9) $$|f(x)| \longrightarrow \infty \text{ as } |x| \longrightarrow \infty.$$

To prove this, the constant term of f is irrelevant, so we may suppose that it is zero. Then $f(x)$ is divisible by x: $f(x) = xg(x)$. By induction on the degree, the assertion is true for $g(x)$, or else $g(x)$ is constant, and it follows for $f(x)$ as well.

Now since $f(x)$ is large for large x, the greatest lower bound m of $|f(x)|$ in the whole complex plane is also the greatest lower bound in a sufficiently large disc $|x| \le r$. Since the disc is compact and $|f(x)|$ is a continuous function, it takes on a minimum value in the disc. □

There are several other proofs of the Fundamental Theorem of Algebra, and one of them is particularly appealing, though it is not as easy to make precise as the one just given. We will present it in outline. As before, our problem is to prove that a nonconstant polynomial

(9.10) $$f(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$$

has a root. If $a_0 = 0$, then 0 is a root, so we may assume that $a_0 \neq 0$. We consider the function $f: \mathbb{C} \longrightarrow \mathbb{C}$ defined by the polynomial (9.10).

Let C_r denote a circle of radius r about the origin. We study the images $f(C_r)$ of the circle C_r. To do this, we use polar coordinates, writing $z = re^{i\theta}$. Then $z^n = r^n e^{in\theta}$. As θ runs from 0 to 2π, the point z winds once around the circle of radius r. At the same time, $n\theta$ runs from 0 to $2\pi n$, so the point z^n winds n times around the circle of radius r^n.

For sufficiently large r, the term z^n is dominant in the expression (9.10), and we will have

$$|f(z) - z^n| \le \tfrac{1}{2}r^n.$$

The proof of this fact is similar to the proof of Lemma (9.6). For our purposes, the factor $\frac{1}{2}$ could be replaced by any positive real number less than 1. This inequality shows us that, as z^n winds n times around the circle of radius r^n, $f(z)$ also winds n times around the origin. A good way to visualize this conclusion is with the dog-on-a-leash model. If someone walks a dog n times around the block, the dog also goes around n times, though following a different path. This will be true provided that the leash is shorter than the radius of the block. Here z^n represents the position of the person at the time θ, and $f(z)$ represents the position of the dog. The length of the leash is $\frac{1}{2}r^n$.

We now vary the radius r. Since f is a continuous function, the image $f(C_r)$ will vary continuously with r. When the radius r is very small, $f(C_r)$ makes a small loop around the constant term a_0 of f. This small loop won't wind around the origin at all. But as we just saw, $f(C_r)$ winds n times around the origin if r is large enough. The only explanation for this is that for some intermediate radius r', $f(C_{r'})$ passes through the origin. This means that for some point α on the circle $C_{r'}$, $f(\alpha) = 0$. This number α is a root of f.

Note that all n loops have to cross the origin, which agrees with the fact that a polynomial of degree n has n roots.

I don't consider this algebra,
but this doesn't mean that algebraists can't do it.

Garrett Birkhoff

EXERCISES

1. Examples of Fields

1. Let F be a field. Find all elements $a \in F$ such that $a = a^{-1}$.
2. Let K be a subfield of $\mathbb{C}$ which is not contained in $\mathbb{R}$. Prove that K is a dense subset of $\mathbb{C}$.
3. Let R be an integral domain containing a field F as subring and which is finite-dimensional when viewed as vector space over F. Prove that R is a field.
4. Let F be a field containing exactly eight elements. Prove or disprove: The characteristic of F is 2.

2. Algebraic and Transcendental Elements

1. Let α be the real cube root of 2. Compute the irreducible polynomial for $1 + \alpha^2$ over $\mathbb{Q}$.
2. Prove Lemma (2.7), that $(1, \alpha, \alpha^2, \ldots, \alpha^{n-1})$ is a basis of $F[\alpha]$.
3. Determine the irreducible polynomial for $\alpha = \sqrt{3} + \sqrt{5}$ over each of the following fields.
   (a) $\mathbb{Q}$ **(b)** $\mathbb{Q}(\sqrt{5})$ **(c)** $\mathbb{Q}(\sqrt{10})$ **(d)** $\mathbb{Q}(\sqrt{15})$

4. Let α be a complex root of the irreducible polynomial $x^3 - 3x + 4$. Find the inverse of $\alpha^2 + \alpha + 1$ in $F(\alpha)$ explicitly, in the form $a + b\alpha + c\alpha^2$, $a, b, c \in \mathbb{Q}$.

5. Let $K = F(\alpha)$, where α is a root of the irreducible polynomial $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$. Determine the element α^{-1} explicitly in terms of α and of the coefficients a_i.

6. Let $\beta = \zeta\sqrt[3]{2}$, where $\zeta = e^{2\pi i/3}$, and let $K = \mathbb{Q}(\beta)$. Prove that -1 can not be written as a sum of squares in K.

3. The Degree of a Field Extension

1. Let F be a field, and let α be an element which generates a field extension of F of degree 5. Prove that α^2 generates the same extension.

2. Let $\zeta = e^{2\pi i/7}$, and let $\eta = e^{2\pi i/5}$. Prove that $\eta \notin \mathbb{Q}(\zeta)$.

3. Define $\zeta_n = e^{2\pi i/n}$. Find the irreducible polynomial over $\mathbb{Q}$ of **(a)** ζ_4, **(b)** ζ_6, **(c)** ζ_8, **(d)** ζ_9, **(e)** ζ_{10}, **(f)** ζ_{12}.

4. Let $\zeta_n = e^{2\pi i/n}$. Determine the irreducible polynomial over $\mathbb{Q}(\zeta_3)$ of **(a)** ζ_6, **(b)** ζ_9, **(c)** ζ_{12}.

5. Prove that an extension K of F of degree 1 is equal to F.

6. Let a be a positive rational number which is not a square in $\mathbb{Q}$. Prove that $\sqrt[4]{a}$ has degree 4 over $\mathbb{Q}$.

7. Decide whether or not i is in the field **(a)** $\mathbb{Q}(\sqrt{-2})$, **(b)** $\mathbb{Q}(\sqrt[4]{-2})$, **(c)** $\mathbb{Q}(\alpha)$, where $\alpha^3 + \alpha + 1 = 0$.

8. Let K be a field generated over F by two elements α, β of relatively prime degrees m, n respectively. Prove that $[K:F] = mn$.

9. Let α, β be complex numbers of degree 3 over $\mathbb{Q}$, and let $K = \mathbb{Q}(\alpha, \beta)$. Determine the possibilities for $[K:\mathbb{Q}]$.

10. Let α, β be complex numbers. Prove that if $\alpha + \beta$ and $\alpha\beta$ are algebraic numbers, then α and β are also algebraic.

11. Let α, β be complex roots of irreducible polynomials $f(x), g(x) \in \mathbb{Q}[x]$. Let $F = \mathbb{Q}[\alpha]$ and $K = \mathbb{Q}[\beta]$. Prove that $f(x)$ is irreducible in K if and only if $g(x)$ is irreducible in F.

12. **(a)** Let $F \subset F' \subset K$ be field extensions. Prove that if $[K:F] = [K:F']$, then $F = F'$. **(b)** Give an example showing that this need not be the case if F is not contained in F'.

13. Let $\alpha_1, \ldots, \alpha_k$ be elements of an extension field K of F, and assume that they are all algebraic over F. Prove that $F(\alpha_1, \ldots, \alpha_k) = F[\alpha_1, \ldots, \alpha_k]$.

14. Prove or disprove: Let α, β be elements which are algebraic over a field F, of degrees d, e respectively. The monomials $\alpha^i\beta^j$ with $i = 0, \ldots, d - 1$, $j = 0, \ldots, e - 1$ form a basis of $F(\alpha, \beta)$ over F.

15. Prove or disprove: Every algebraic extension is a finite extension.

4. Constructions with Ruler and Compass

1. Express $\cos 15°$ in terms of square roots.

2. Prove that the regular pentagon can be constructed by ruler and compass **(a)** by field theory, and **(b)** by finding an explicit construction.

3. Derive formula (4.12).
4. Determine whether or not the regular 9-gon is constructible by ruler and compass.
5. Is it possible to construct a square whose area is equal to that of a given triangle?
6. Let α be a real root of the polynomial $x^3 + 3x + 1$. Prove that α can not be constructed by ruler and compass.
7. Given that π is a transcendental number, prove the impossibility of squaring the circle by ruler and compass. (This means constructing a square whose area is the same as the area of a circle of unit radius.)
8. Prove the impossibility of "duplicating the cube," that is, of constructing the side length of a cube whose volume is 2.
9. **(a)** Referring to the proof of Proposition (4.8), prove that the discriminant D is negative if and only if the circles do not intersect.
   (b) Determine the line which appears at the end of the proof of Proposition (4.8) geometrically if $D \geq 0$ and also if $D < 0$.
10. Prove that if a prime integer p has the form $2^r + 1$, then it actually has the form $2^{2^k} + 1$.
11. Let C denote the field of constructible real numbers. Prove that C is the smallest subfield of $\mathbb{R}$ with the property that if $a \in C$ and $a > 0$, then $\sqrt{a} \in C$.
12. The points in the plane can be considered as complex numbers. Describe the set of constructible points explicitly as a subset of $\mathbb{C}$.
13. Characterize the constructible real numbers in the case that three points are given in the plane to start with.

*14. Let the rule for construction in three-dimensional space be as follows:
   (i) Three non-collinear points are given. They are considered to be constructed.
   (ii) One may construct a plane through three non-collinear constructed points.
   (iii) One may construct a sphere with center at a constructed point and passing through another constructed point.
   (iv) Points of intersection of constructed planes and spheres are considered to be constructed if they are isolated points, that is, if they are not part of an intersection curve.

   Prove that one can introduce coordinates, and characterize the coordinates of the constructible points.

5. *Symbolic Adjunction of Roots*

1. Let F be a field of characteristic zero, let f' denote the derivative of a polynomial $f \in F[x]$, and let g be an irreducible polynomial which is a common divisor of f and f'. Prove that g^2 divides f.
2. For which fields F and which primes p does $x^p - x$ have a multiple root?
3. Let F be a field of characteristic p.
   (a) Apply (5.7) to the polynomial $x^p + 1$.
   (b) Factor this polynomial into irreducible factors in $F[x]$.
4. Let $\alpha_1, \ldots, \alpha_n$ be the roots of a polynomial $f \in F[x]$ of degree n in an extension field K. Find the best upper bound that you can for $[F(\alpha_1, \ldots, \alpha_n) : F]$.

6. Finite Fields

1. Identify the group $\mathbb{F}_4^+$.
2. Write out the addition and multiplication tables for $\mathbb{F}_4$ and for $\mathbb{Z}/(4)$, and compare them.
3. Find a thirteenth root of 3 in the field $\mathbb{F}_{13}$.
4. Determine the irreducible polynomial over $\mathbb{F}_2$ for each of the elements (6.12) of $\mathbb{F}_8$.
5. Determine the number of irreducible polynomials of degree 3 over the field $\mathbb{F}_3$.
6. **(a)** Verify that (6.9, 6.10, 6.13) are irreducible factorizations over $\mathbb{F}_2$.
   (b) Verify that (6.11, 6.13) are irreducible factorizations over $\mathbb{Z}$.
7. Factor $x^9 - x$ and $x^{27} - x$ in $\mathbb{F}_3$. Prove that your factorizations are irreducible.
8. Factor the polynomial $x^{16} - x$ in the fields **(a)** $\mathbb{F}_4$ and **(b)** $\mathbb{F}_8$.
9. Determine all polynomials $f(x)$ in $\mathbb{F}_q[x]$ such that $f(\alpha) = 0$ for all $\alpha \in \mathbb{F}_q$.
10. Let K be a finite field. Prove that the product of the nonzero elements of K is -1.
11. Prove that every element of $\mathbb{F}_p$ has exactly one pth root.
12. Complete the proof of Proposition (6.19) by showing that the difference $\alpha - \beta$ of two roots of $x^q - x$ is a root of the same polynomial.
13. Let p be a prime. Describe the integers n such that there exist a finite field K of order n and an element $\alpha \in K^\times$ whose order in $K^\times$ is p.
14. Work this problem without appealing to Theorem (6.4).
    (a) Let $F = \mathbb{F}_p$. Determine the number of monic irreducible polynomials of degree 2 in $F[x]$.
    (b) Let $f(x)$ be one of the polynomials described in (a). Prove that $K = F[x]/(f)$ is a field containing p^2 elements and that the elements of K have the form $a + b\alpha$, where $a, b \in F$ and α is a root of f in K. Show that every such element $a + b\alpha$ with $b \neq 0$ is the root of an irreducible quadratic polynomial in $F[x]$.
    (c) Show that every polynomial of degree 2 in $F[x]$ has a root in K.
    (d) Show that all the fields K constructed as above for a given prime p are isomorphic.
15. The polynomials $f(x) = x^3 + x + 1$, $g(x) = x^3 + x^2 + 1$ are irreducible over $\mathbb{F}_2$. Let K be the field extension obtained by adjoining a root of f, and let L be the extension obtained by adjoining a root of g. Describe explicitly an isomorphism from K to L.
16. **(a)** Prove Lemma (6.21) for the case $F = \mathbb{C}$ by looking at the roots of the two polynomials.
    (b) Use the principle of permanence of identities to derive the conclusion when F is an arbitrary ring.

7. Function Fields

1. Determine a real polynomial in three variables whose locus of zeros is the projected Riemann surface (7.9).
2. Prove that the set $\mathcal{F}(U)$ of continuous functions on U' forms a ring.
3. Let $f(x)$ be a polynomial in $F[x]$, where F is a field. Prove that if there is a rational function $r(x)$ such that $r^2 = f$, then r is a polynomial.
4. Referring to the proof of Proposition (7.11), explain why the map $F \longrightarrow \mathcal{F}(S)$ defined by $g(x) \rightsquigarrow g(X)$ is a homomorphism.

5. Determine the branch points and the gluing data for the Riemann surfaces of the following polynomials.
(**a**) $y^2 - x^2 + 1$ (**b**) $y^5 - x$ (**c**) $y^4 - x - 1$ (**d**) $y^3 - xy - x$
(**e**) $y^3 - y^2 - x$ (**f**) $y^3 - x(x-1)$ (**g**) $y^3 - x(x-1)^2$ (**h**) $y^3 + xy^2 + x$
(**i**) $x^2y^2 - xy - x$

6. (**a**) Determine the number of isomorphism classes of function fields K of degree 3 over $F = \mathbb{C}(x)$ which are ramified only at the points ± 1.
(**b**) Describe the gluing data for the Riemann surface corresponding to each isomorphism class of fields as a pair of permutations.
(**c**) For each isomorphism class, determine a polynomial $f(x, y)$ such that $K = F[x]/(f)$ represents the isomorphism class.

***7.** Prove the Riemann Existence Theorem for quadratic extensions.

***8.** Let S be a branched covering constructed with branch points $\alpha_1, \dots, \alpha_r$, curves $C_1, \dots, C_r$, and permutations $\sigma_1, \dots, \sigma_r$. Prove that S is connected if and only if the subgroup Σ of the symmetric group S_n which is generated by the permutations σ_ν operates transitively on the indices $\mathbf{1}, \dots, \mathbf{n}$.

***9.** It can be shown that the Riemann surface S of a function field is homeomorphic to the complement of a finite set of points in a compact oriented two-dimensional manifold $\overline{S}$. The *genus* of such a surface is defined to be the number of holes in the corresponding manifold $\overline{S}$. So if $\overline{S}$ is a sphere, the genus of S is 0, while if $\overline{S}$ is a torus, the genus of S is 1. The genus of a function field is defined to be the genus of its Riemann surface. Determine the genus of the field defined by each polynomial.
(**a**) $y^2 - (x^2-1)(x^2-4)$ (**b**) $y^2 - x(x^2-1)(x^2-4)$ (**c**) $y^3 + y + x$
(**d**) $y^3 - x(x-1)$ (**e**) $y^3 - x(x-1)^2$

8. Transcendental Extensions

1. Let $K = F(\alpha)$ be a field extension generated by an element α, and let $\beta \in K$, $\beta \notin F$. Prove that α is algebraic over the field $F(\beta)$.

2. Prove that the isomorphism $\mathbb{Q}(\pi) \longrightarrow \mathbb{Q}(e)$ sending $\pi \rightsquigarrow e$ is discontinuous.

3. Let $F \subset K \subset L$ be fields. Prove that $\text{tr deg}_F L = \text{tr deg}_F K + \text{tr deg}_K L$.

4. Let $(\alpha_1, \dots, \alpha_n) \subset K$ be an algebraically independent set over F. Prove that an element $\beta \in K$ is transcendental over $F(\alpha_1, \dots, \alpha_n)$ if and only if $(\alpha_1, \dots, \alpha_n; \beta)$ is algebraically independent.

5. Prove Theorem (8.3).

9. Algebraically Closed Fields

1. Derive Corollary (9.5) from Theorem (9.4).

2. Prove that the field $\overline{F}$ constructed in this text as the union of finite fields is algebraically closed.

***3.** With notation as at the end of the section, a comparison of the images $f(C_r)$ for varying radii shows another interesting geometric feature: For large r, the curve $f(C_r)$ has n loops. This can be expressed formally by saying that its total curvature is $2\pi n$. For small r, the linear term $a_1 z + a_0$ dominates $f(z)$. Then $f(C_r)$ makes a single loop around a_0. Its

total curvature is only 2π. Something happens to the loops and the curvature, as r varies. Explain.

*4. If you have access to a computer with a good graphics system, use it to illustrate the variation of $f(C_r)$ with r. Use log-polar coordinates $(\log r, \theta)$.

Miscellaneous Exercises

1. Let $f(x)$ be an irreducible polynomial of degree 6 over a field F, and let K be a quadratic extension of F. Prove or disprove: Either f is irreducible over K, or else f is a product of two irreducible cubic polynomials over K.
2. (a) Let p be an odd prime. Prove that exactly half of the elements of $\mathbb{F}_p^{\times}$ are squares and that if α, β are nonsquares, then $\alpha\beta$ is a square.
   (b) Prove the same as (a) for any finite field of odd order.
   (c) Prove that in a finite field of even order, every element is a square.
3. Write down the irreducible polynomial for $\alpha = \sqrt{2} + \sqrt{3}$ over $\mathbb{Q}$ and prove that it is reducible modulo p for every prime p.

*4. (a) Prove that any element of $GL_2(\mathbb{Z})$ of finite order has order $1, 2, 3, 4$, or 6.
   (b) Extend this theorem to $GL_3(\mathbb{Z})$, and show that it fails in $GL_4(\mathbb{Z})$.

5. Let c be a real number, not ± 2. The plane curve C: $x^2 + cxy + y^2 = 1$ can be parametrized rationally. To do this, we choose the point $(0, 1)$ on C and parametrize the lines through this point by their slope: $L_t: y = tx + 1$. The point at which the line L_t intersects C can be found algebraically.
   (a) Find the equation of this point explicitly.
   (b) Use this procedure to find all solutions of the equation $x^2 + cxy + y^2 = 1$ in the field $F = \mathbb{F}_p$, when c is in that field and $c \neq \pm 2$.
   (c) Show that the number of solutions is $p - 1$, p, or $p + 1$, and describe how this number depends on the roots of the polynomial $t^2 + ct + 1$.
6. The *degree* of a rational function $f(x) = p(x)/q(x) \in \mathbb{C}(x)$ is defined to be the maximum of the degrees of p and q, when p, q are chosen to be relatively prime. Every rational function f defines a map $P' \longrightarrow P'$, by $x \rightsquigarrow f(x)$. We will denote this map by f too.
   (a) Suppose that f has degree d. Show that for any point y_0 in the plane, the fibre $f^{-1}(y_0)$ contains at most d points.
   (b) Show that $f^{-1}(y_0)$ consists of precisely d points, except for a finite number of y_0. Identify the values y_0 where there are fewer than d points in terms of f and df/dx.

*7. (a) Prove that a rational function $f(x)$ generates the field of rational functions $\mathbb{C}(x)$ if and only if it is of the form $(ax + b)/(cx + d)$, with $ad - bc \neq 0$.
   (b) Identify the group of automorphisms of $\mathbb{C}(x)$ which are the identity on $\mathbb{C}$.

*8. Let K/F be an extension of degree 2 of rational function fields, say $K = \mathbb{C}(t)$ and $F = \mathbb{C}(x)$. Prove that there are generators x', t' for the two fields, such that $t = (\alpha t' + \beta)/(\gamma t' + \delta)$ and $x = (ax' + b)/(cx' + d)$, $\alpha, \beta, \gamma, \delta, a, b, c, d \in \mathbb{C}$, such that $t'^2 = x'$.

*9. Fill in the following outline to give an algebraic proof of the fact that $K = \mathbb{C}(x)[y]/(y^2 - x^3 + x)$ is not a pure transcendental extension of $\mathbb{C}$. Suppose that $K = \mathbb{C}(t)$ for some t. Then x and y are rational functions of t.

(a) Using the result of the previous problem and replacing t by t' as necessary, reduce to the case that $x = (at^2 + b)/(ct^2 + d)$.

(b) Say that $y = p(t)/q(t)$. Then the equation $y^2 = x(x + 1)(x - 1)$ reads

$$\frac{p(t)^2}{q(t)^2} = \frac{(at^2 + b)((a + c)t^2 + b + d)((a - c)t^2 + b - d)}{(ct^2 + d)^3}.$$

Either the numerators and denominators on the two sides agree, or else there is cancellation on the right side.

(c) Complete the proof by analyzing the two possibilities given in (b).

***10.** (a) Prove that the homomorphism $SL_2(\mathbb{Z}) \longrightarrow SL_2(\mathbb{F}_p)$ obtained by reducing the matrix entries modulo 2 is surjective.

(b) Prove the analogous assertion for SL_n.

***11.** Determine the conjugacy classes of elements order 2 in $GL_n(\mathbb{Z})$.

Chapter 14

Galois Theory

En un mot les calculs sont impraticables.

Evariste Galois

1. THE MAIN THEOREM OF GALOIS THEORY

In the last chapter we studied algebraic field extensions, using extensions generated by a single element as the basic tool. This amounts to studying the properties of a single root of an irreducible polynomial

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0. \tag{1.1}$$

Galois theory, the topic of this chapter, is the theory of *all* the roots of such a polynomial and of the symmetries among them.

We will restrict our attention to fields of *characteristic zero* in this chapter. It is to be understood that all fields occurring have characteristic zero, and we will not mention this assumption explicitly from now on.

The notation K/F will indicate that K is an extension field of F. This notation is traditional, though there is some danger of confusion with the notation R/I for the quotient of a ring R by an ideal I.

As we have seen, computation in a field $F(\alpha)$ generated by a single root can easily be made by identifying it with the formally constructed field $F[x]/(f)$. But suppose that an irreducible polynomial $f(x)$ factors into linear factors in a field extension K, and that its roots in K are $\alpha_1,\ldots,\alpha_n$. How to compute with all these roots at the same time isn't clear. To do so we have to know how the roots are related, and this depends on the particular case. In principle, the relations can be obtained by expanding the equation $f(x) = (x - \alpha_1)(x - \alpha_2)\cdots(x - \alpha_n)$. Doing so, we find that the sum of the roots is $-a_{n-1}$, that their product is $\pm a_0$, and so on. However, it may not be easy to interpret these relations directly.

The fundamental discovery which arose through the work of several people, especially of Lagrange and Galois, is that the relationships between the roots can be understood in terms of symmetry. The original model for this symmetry is complex conjugation, which permutes the two roots $\pm i$ of the irreducible real polynomial $x^2 + 1$, while leaving the real numbers fixed. We will begin by observing that such a symmetry exists for any quadratic field extension.

An extension K/F of degree 2 is generated by any element α of K which is not in F. Moreover, α is a root of an irreducible quadratic polynomial

$$f(x) = x^2 + bx + c \tag{1.2}$$

with coefficients in F. Then $\alpha' = -b - \alpha$ is also a root of f, so this polynomial splits into linear factors over K: $f(x) = (x - \alpha)(x - \alpha')$.

The fact that α and α' are roots of the same irreducible polynomial provides us with our symmetry. According to Proposition (2.9) of Chapter 13, there is an isomorphism

$$\sigma: F(\alpha) \longrightarrow F(\alpha'), \tag{1.3}$$

which is the identity on F and which sends $\alpha \rightsquigarrow \alpha'$. But either root generates the extension: $F(\alpha) = K = F(\alpha')$. Therefore σ is an automorphism of K.

This automorphism switches the two roots α, α'. For, since σ is the identity on F, it fixes b, and $\alpha + \alpha' = b$. So if $\sigma(\alpha) = \alpha'$, we must have $\sigma(\alpha') = \alpha$. It follows that σ^2 sends $\alpha \rightsquigarrow \alpha$ and, since α generates K over F, that σ^2 is the identity.

Note also that σ is not the identity automorphism, because the two roots α, α' are distinct. If α were a double root of the quadratic polynomial (1.2), the quadratic formula would give $\alpha = -\frac{1}{2}b$. This would imply $\alpha \in F$, contrary to our hypothesis that f is irreducible.

Since our field F is assumed to have characteristic zero, the quadratic extension K can be obtained by adjoining a square root δ of the discriminant $D = b^2 - 4c$, a root of the irreducible polynomial $x^2 - D$. Its other root is $-\delta$, and σ interchanges the two square roots.

Whenever K is obtained by adjoining a square root δ, there is an automorphism which sends $\delta \rightsquigarrow -\delta$. For example, let $\alpha = 1 + \sqrt{2}$, and let $K = \mathbb{Q}(\alpha)$. The irreducible polynomial for α over $\mathbb{Q}$ is $x^2 - 2x - 1$, and the other root of this polynomial is $\alpha' = 1 - \sqrt{2}$. There is an automorphism σ of K which sends $\sqrt{2} \rightsquigarrow -\sqrt{2}$ and $\alpha \rightsquigarrow \alpha'$. It is important to note right away that such an automorphism will *not* be continuous when K is considered as a subfield of $\mathbb{R}$. It is a symmetry of the algebraic structure of K, but it does not respect the geometry given by the embedding of K into the real line.

By definition, an *F-automorphism* of an extension field K is an automorphism which is the identity on the subfield F [see Chapter 13 (2.10)]. In other words, an automorphism σ of K is an F-automorphism if $\sigma(c) = c$ for all $c \in F$. Thus complex conjugation is an $\mathbb{R}$-automorphism of $\mathbb{C}$, and the symmetry σ we have just

found is an F-automorphism of the quadratic extension K. It is not difficult to show that σ is the only F-automorphism of this extension other than the identity.

The group of all F-automorphisms of K is called the *Galois group* of the field extension. We often denote this group by $G(K/F)$. When K/F is a quadratic extension, the Galois group $G(K/F)$ is a group of order 2.

Let us now consider the next simplest example, that of a biquadratic extension. We will call a field extension K/F *biquadratic* if $[K:F] = 4$ and if K is generated by the roots of *two* irreducible quadratic polynomials. Every such extension has the form

$$K = F(\alpha, \beta), \tag{1.4}$$

where $\alpha^2 = a$ and $\beta^2 = b$, and where a, b are elements of F. The element β generates an intermediate field—a field $F(\beta)$ between F and K. Since $K = F(\alpha, \beta)$, the requirement that $[K:F] = 4$ implies that $F(\beta)$ has degree 2 over F and that α is not in the field $F(\beta)$. So the polynomial $x^2 - a$ is irreducible over $F(\beta)$. Similarly, the polynomial $x^2 - b$ is irreducible over the intermediate field $F(\alpha)$.

Notice that K is an extension of $F(\beta)$ of degree 2, generated by α. Let us apply what we have just learned about quadratic extensions to this extension. Substituting $F(\beta)$ for F, we find that there is an $F(\beta)$-automorphism of K which interchanges the two roots $\pm\alpha$ of $x^2 - a$. Call this automorphism σ. Since it is the identity on $F(\beta)$, σ is also the identity on F, so it is an F-automorphism too. Similarly, there is an $F(\alpha)$-automorphism τ of K which interchanges the roots $\pm\beta$ of $x^2 - b$, and τ is also an F-automorphism.

The two automorphisms we have found operate on the roots α, β as follows:

$$\begin{array}{ll} \alpha \overset{\sigma}{\rightsquigarrow} -\alpha & \alpha \overset{\tau}{\rightsquigarrow} \alpha \\ \beta \overset{\sigma}{\rightsquigarrow} \beta & \beta \overset{\tau}{\rightsquigarrow} -\beta. \end{array} \tag{1.5}$$

Composing these operations, we find that $\sigma\tau$ changes the signs of both roots α, β and that the automorphisms σ^2, τ^2, and $\sigma\tau\sigma\tau$ leave α and β fixed. Since K is generated over F by the roots, these last three automorphisms are all equal to the identity. Therefore the four automorphisms $\{1, \sigma, \tau, \sigma\tau\}$ form a group of order 4, with relations

$$\sigma^2 = 1, \quad \tau^2 = 1, \quad \sigma\tau = \tau\sigma.$$

We have shown that the Galois group $G(K/F)$ contains the Klein four group. In fact it is equal to that group, as we shall see in a moment.

For example, let $F = \mathbb{Q}$, $\alpha = i$, and $\beta = \sqrt{2}$, so that $K = \mathbb{Q}(i, \sqrt{2})$. In this case, the automorphism σ is complex conjugation, while τ sends $\sqrt{2} \rightsquigarrow -\sqrt{2}$, fixing i.

For quadratic or biquadratic extensions, the degree $[K : F]$ is equal to the order of the Galois group $G(K/F)$. We will now state two theorems, Theorems (1.6) and (1.11), which describe the general circumstances under which this happens. These theorems will be proved in later sections of the chapter.

(1.6) **Theorem.** For any finite extension K/F, the order $|G(K/F)|$ of the Galois group divides the degree $[K : F]$ of the extension.

A finite field extension K/F is called a *Galois extension* if the order of the Galois group is equal to the degree:

$$|G(K/F)| = [K : F]. \tag{1.7}$$

Theorem (1.6) shows that the Galois group of a biquadratic extension has order at most 4. Since we already have four automorphisms in hand, there are no others, and the Galois group is the Klein four group, as was asserted. All quadratic and biquadratic extensions are Galois.

If G is a group of automorphisms of a field K, the set of elements of K which are fixed by all the automorphisms in G forms a subfield, called the *fixed field* of G. The fixed field is often denoted by K^G:

$$K^G = \{\alpha \in K \mid \varphi(\alpha) = \alpha \text{ for all } \varphi \in G\}. \tag{1.8}$$

One consequence of Theorem (1.6) is that when K/F is a Galois extension, the only elements of K which are fixed by the whole Galois group are the elements of F:

(1.9) **Corollary.** Let K/F be a Galois extension, with Galois group $G = G(K/F)$. The fixed field of G is F.

For let L denote the fixed field. Then $F \subset L$, and this inclusion shows that every L-automorphism of K is also an F-automorphism, that is, that $G(K/L) \subset G$. On the other hand, by definition of the fixed field, every element of G is an L-automorphism. So $G(K/L) = G$. Now $|G| = [K : F]$ because K/F is a Galois extension, and by Theorem (1.6), $|G|$ divides $[K : L]$. Since $F \subset L \subset K$, this shows that $[K : F] = [K : L]$, hence that $F = L$. □

This corollary is important because it provides a method for checking that an element of a Galois extension K is actually in the field F. We will use it frequently.

Being Galois is a strong restriction on a field extension, but nevertheless there are many Galois extensions. This is the key fact which led to Galois' theory. In order to state the theorem which describes the Galois extensions, we need one more definition.

(1.10) **Definition.** Let $f(x) \in F[x]$ be a nonconstant monic polynomial. A *splitting field* for $f(x)$ over F is an extension field K of F such that

(i) $f(x)$ factors into linear factors in K: $f(x) = (x-\alpha_1)\cdots(x-\alpha_n)$, with $\alpha_i \in K$;
(ii) K is generated by the roots of $f(x)$: $K = F(\alpha_1,\ldots,\alpha_n)$.

The second condition just says that K is the smallest extension of F which contains all the roots. The biquadratic extension (1.4) is a splitting field of the polynomial $f(x) = (x^2 - a)(x^2 - b)$.

Every polynomial $f(x) \in F[x]$ has a splitting field. To find one, we choose a field extension L in which f splits into linear factors [Chapter 13 (5.3)] and then take for K the subfield $F(\alpha_1,\dots,\alpha_n)$ of L generated by the roots.

(1.11) **Theorem.** If K is a splitting field of a polynomial $f(x)$ over F, then K is a Galois extension of F. Conversely, every Galois extension is a splitting field of some polynomial $f(x) \in F[x]$.

(1.12) **Corollary.** Every finite extension is contained in a Galois extension.

To derive this corollary from the theorem, let K/F be a finite extension, let $\alpha_1,\dots,\alpha_n$ be generators for K over F, and let $f_i(x)$ be the monic irreducible polynomial for α_i over F. We extend K to a splitting field L of the product $f = f_1 \cdots f_n$ over K. Then L will also be a splitting field of f over F. So L is the required Galois extension. □

(1.13) **Corollary.** Let K/F be a Galois extension, and let L be an intermediate field: $F \subset L \subset K$. Then K/L is a Galois extension too.

For, if K is the splitting field of a polynomial $f(x)$ over F, then it is also the splitting field of the same polynomial over the larger field L, so K is a Galois extension of L. □

Let us go back to biquadratic extensions. We can prove that the Galois group of such an extension has order 4 without appealing to Theorem (1.6). All that is needed is the following elementary proposition:

(1.14) **Proposition.**

(a) Let K be an extension of a field F, let $f(x)$ be a polynomial with coefficients in F, and let σ be an F-automorphism of K. If α is a root of $f(x)$ in K, then $\sigma(\alpha)$ is also a root.

(b) Let K be a field extension generated over F by elements $\alpha_1,\dots,\alpha_r$, and let σ be an F-automorphism of K. If σ fixes each of the generators α_i, then σ is the identity automorphism.

(c) Let K be a splitting field of a polynomial $f(x)$ over F. The Galois group $G(K/F)$ operates faithfully on the set $\{\alpha_1,\dots,\alpha_n\}$.

Proof. Part (a) was proved in the last chapter [Chapter 13 (2.10)]. To prove part (b), assume that K is generated by $\alpha_1,\dots,\alpha_n$. Then every element of K can be expressed as a polynomial in $\alpha_1,\dots,\alpha_n$ with coefficients in F [Chapter 13 (2.6b)]. If σ is an automorphism which is the identity on F and which also fixes each of the elements α_i, then it fixes every polynomial in $\{\alpha_i\}$ with coefficients in F; hence it is the identity. The third assertion (c) follows from the first two: The first tells us that every $\sigma \in G(K/F)$ permutes the set $\{\alpha_1,\dots,\alpha_n\}$, and the second tells us that the operation on this set is faithful. □

Proposition (1.14) does not address the most interesting question: *Which permutations of the roots of a polynomial extend to automorphisms of the splitting field?* This question is the central theme of Galois theory.

Let us apply Proposition (1.14) to the biquadratic extension (1.4). Part (a), applied to the polynomial $x^2 - a$, shows that any F-automorphism φ of K permutes the roots $\pm\alpha$. Similarly, φ permutes $\pm\beta$. Only four permutations of $\{\pm\alpha, \pm\beta\}$ act in this way. Since the elements α, β generate K, (1.14b) tells us that an F-automorphism which fixes both of them is the identity. So the four automorphisms which we have already found are the only ones. This proves that $G(K/F)$ is the Klein four group.

One of the most important parts of Galois theory is the determination of the *intermediate fields* L, those sandwiched between F and $K : F \subset L \subset K$. The Main Theorem of Galois theory asserts that when K/F is a Galois extension, the intermediate fields are in bijective correspondence with the subgroups of the Galois group. The importance of this correspondence is not immediately clear. We will have to see it used to understand it.

The intermediate field corresponding to a subgroup H of $G(K/F)$ is the fixed field K^H of H, which was defined above. In the other direction, if L is an intermediate field, the Galois group $G(K/L)$ is a subgroup of $G(K/F)$. This is the subgroup which corresponds to L.

(1.15) **Theorem.** *The Main Theorem:* Let K be a Galois extension of a field F, and let $G = G(K/F)$ be its Galois group. The function

$$H \rightsquigarrow K^H$$

is a bijective map from the set of subgroups of G to the set of intermediate fields $F \subset L \subset K$. Its inverse function is

$$L \rightsquigarrow G(K/L).$$

This correspondence has the property that if $H = G(K/L)$, then

$$(1.16) \qquad [K : L] = |H|, \quad \text{hence} \quad [L : F] = [G : H].$$

We will prove this theorem in Section 5.

The fields F and K are included among the intermediate fields. The subgroup which corresponds to the field F is the whole group G [see (1.9)], and the one corresponding to K is the trivial subgroup $\{1\}$.

Let us go back to our example of the biquadratic extension $K = \mathbb{Q}(i, \sqrt{2})$, for which σ is complex conjugation, while τ interchanges $\sqrt{2} \rightsquigarrow -\sqrt{2}$. Its Galois group, the Klein four group, has three proper subgroups:

$$H_1 = \{1, \sigma\}, \quad H_2 = \{1, \tau\}, \quad H_3 = \{1, \sigma\tau\}.$$

According to the Main Theorem, there are three proper intermediate fields, namely the fixed fields L_i of these subgroups. They are easily determined:

$$L_1 = \mathbb{Q}(\sqrt{2}), \quad L_2 = \mathbb{Q}(i), \quad \text{and} \quad L_3 = \mathbb{Q}(i\sqrt{2}).$$

A Galois group is finite, so it has finitely many subgroups. But without the Main Theorem, it isn't obvious that there are only finitely many intermediate fields. It might seem natural to expect two randomly chosen elements of a Galois extension K/F to generate different subfields. This tends not to happen, and in fact most elements will generate the whole extension K. The case of the biquadratic extension $K = \mathbb{Q}(i, \sqrt{2})$ will illustrate this point. Let γ be any element of K. The field $\mathbb{Q}(\gamma)$ generated by γ must be one of the intermediate fields we have found. So if γ is not contained in $\mathbb{Q}(i)$, $\mathbb{Q}(\sqrt{2})$, or $\mathbb{Q}(i\sqrt{2})$, then $\mathbb{Q}(\gamma) = K$. Now the set $(1, i, \sqrt{2}, i\sqrt{2})$ is a basis for K over F, so we may write an arbitrary element γ in the form

$$\gamma = c_1 + c_2 i + c_3\sqrt{2} + c_4 i\sqrt{2}, \quad \text{with } c_i \in \mathbb{Q}.$$

This element is not in one of the three proper intermediate fields unless two of the coefficients c_2, c_3, c_4 are zero. The element $i + \sqrt{2}$, for example, generates the whole extension K. We will return to this point in Section 4.

2. CUBIC EQUATIONS

Having examined biquadratic extensions in the last section, we now turn to the next general class of examples, the splitting fields of cubic polynomials. Cubic equations

$$(2.1) \qquad f(x) = x^3 + a_2x^2 + a_1x + a_0 = 0$$

were solved explicitly in terms of square roots and cube roots in the sixteenth century by the mathematicians Tartaglia and Cardano. We will begin by reviewing their *remarkable ad hoc solution.*

The computation is simpler when the coefficient of degree 2 in $f(x)$ vanishes. The quadratic term in our general equation (2.1) can be eliminated by the substitution

$$(2.2) \qquad x = x_1 - a_2/3.$$

Let us write a cubic whose quadratic term vanishes as

$$(2.3) \qquad f(x) = x^3 + px + q,$$

where the coefficients p, q are elements of the field F. Cardano's solution of the equation $f = 0$ starts with the substitution $x = u - v$. Collecting terms in $f(u - v)$, we find

$$f(u - v) = (u^3 - v^3) - (3uv - p)(u - v) + q.$$

The point of replacing the variable x by a sum of variables is that we can now split our equation apart. Clearly, $f(u - v) = 0$ if the two equations

$$3uv - p = 0, \quad u^3 - v^3 + q = 0$$

hold. And since we have two variables, we may hope to obtain solutions to such a pair of equations, though it isn't clear a priori that this will help. We solve the first

equation for $v = p/3u$ and substitute into the second. Clearing the denominator gives

$$3^3u^6 - p^3 + 3^3u^3q = 0.$$

Miraculously, this equation is quadratic in u^3. Setting $y = u^3$, it reduces to

$$3^3y^2 + 3^3qy - p^3 = 0. \tag{2.4}$$

This equation can be solved by the quadratic formula:

$$y = -\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}. \tag{2.5}$$

Thus we obtain *Cardano's Formula* $x = u - v$, where

(2.6)
$$u = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}, \quad v = \sqrt[3]{u^3 + q} = \sqrt[3]{+\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}.$$

We will be able to prove the existence of a solution of this general type later, without explicit computation [see (7.6)].

Let us now examine the Galois theory of an irreducible cubic polynomial $f(x)$. We may assume that $f(x)$ has the form (2.3). Let K be a splitting field of $f(x)$ over F, and let $\alpha_1, \alpha_2, \alpha_3$ be the three roots of $f(x)$ in K, ordered in an arbitrary way, so that

$$f(x) = x^3 + px + q = (x-\alpha_1)(x-\alpha_2)(x-\alpha_3). \tag{2.7}$$

Expanding the right side of this equation, we obtain the relations

$$\begin{aligned} \alpha_1 + \alpha_2 + \alpha_3 &= 0 \\ \alpha_1\alpha_2 + \alpha_2\alpha_3 + \alpha_1\alpha_3 &= p \\ \alpha_1\alpha_2\alpha_3 &= -q. \end{aligned} \tag{2.8}$$

The first of these relations shows that the third root α_3 is in the field generated by the first two roots. Thus we have a chain of fields

$$F \subset F(\alpha_1) \subset K,$$

and $K = F(\alpha_1, \alpha_2) = F(\alpha_1, \alpha_2, \alpha_3)$. Let us denote $F(\alpha_1)$ by L. There are two fundamentally different cases which may arise, namely either

$$L = K \quad \text{or} \quad L < K. \tag{2.9}$$

In terms of the roots, the first case occurs when the last two roots α_2 and α_3 can be expressed in terms of α_1 and elements of F, that is, if they can be written as polynomials in α_1 with coefficients in F [see Chapter 13 (2.6)]. The second case occurs when the last two roots can not be expressed in this way.

For example, let $f(x) = x^3 - 2$. The three roots of this polynomial are $\alpha_1 = \sqrt[3]{2}$, $\alpha_2 = \zeta\sqrt[3]{2}$, $\alpha_3 = \zeta^2\sqrt[3]{2}$, where $\sqrt[3]{2}$ denotes the real cube root of 2 and $\zeta = e^{2\pi i/3}$. Since α_1 is real, the field $\mathbb{Q}(\alpha_1)$ is contained in $\mathbb{R}$. It doesn't contain the

other two roots, which are complex. Hence if $F = \mathbb{Q}$ and $L = \mathbb{Q}(\alpha_1)$, we are in the second case. On the other hand, if we let $F = \mathbb{Q}(\zeta)$, then $F(\alpha_1)$ contains α_2, so we are in the first case.

To analyze the dichotomy (2.9), we consider the way the irreducible polynomial $f(x)$ factors in the field L. By assumption, $f(x)$ is irreducible in $F[x]$, and it factors into linear factors in $K[x]$. In the ring $L[x]$, $f(x)$ has the factor $(x - \alpha_1)$:

(2.10) $$f(x) = (x - \alpha_1)h(x),$$

where $h(x)$ is a quadratic polynomial with coefficients in L. Division by $x - \alpha_1$ gives the same result if it is carried out in the larger field K. Looking at (2.7), we see that $h(x) = (x - \alpha_2)(x - \alpha_3)$ in $K[x]$. Therefore $L < K$ if and only if $h(x)$ is irreducible over L. In this case, the degree of $L(\alpha_2) = K$ over L is 2. Also, since we assume $f(x)$ irreducible over F, $[L : F] = 3$ in either case. So we have

(2.11) $$[K : F] = \begin{cases} 3 \text{ if } L = K \\ 6 \text{ if } L < K \end{cases}.$$

(2.12) **Example.** The polynomial $f(x) = x^3 + 3x + 1$ is irreducible over $\mathbb{Q}$, and it has only one real root. To see that there is only one real root, we note that the derivative of f does not vanish on the real line. Therefore $f(x)$ defines an increasing function of the real variable x. It takes the value 0 only once. The real root does not generate the splitting field K, which also contains two complex roots. So $[K : \mathbb{Q}] = 6$ in this case.

On the other hand, the splitting field of the polynomial $f(x) = x^3 - 3x + 1$ over $\mathbb{Q}$ has degree 3. One of its roots is $\eta_1 = 2 \cos 2\pi/9 = \zeta + \zeta^8$, where $\zeta = e^{2\pi i/9}$. Having the polynomial in hand, we can check this directly. But actually, we made this example by computing the irreducible polynomial for η_1 over $\mathbb{Q}$. The way to compute this polynomial is to guess its other roots. We note that η_1 is the sum of a ninth root of 1 and its inverse. There are two other sums of this sort: $\eta_2 = \zeta^2 + \zeta^7$ and $\eta_3 = \zeta^4 + \zeta^5$. We guess that these are the other roots and expand $(x - \eta_1)(x - \eta_2)(x - \eta_3)$, obtaining f. In this example, η_2 happens to be equal to $\eta_1{}^2 - 2$, and $\eta_3 = -\eta_1 - \eta_2$. So $K = F(\eta_1)$. □

We go back to a general cubic equation. According to Theorem (1.11), the order of the Galois group $G = G(K/F)$ is the degree of the field extension $[K : F]$. For cubic equations, this degree determines the group G completely. Namely, Proposition (1.14) tells us that G operates faithfully on the set $\{\alpha_1, \alpha_2, \alpha_3\}$ of roots. These roots are distinct [Chapter 13 (5.8)]. So G is a subgroup of the symmetric group S_3, which has order 6. If $[K : F] = 6$, then G is the whole symmetric group. In this case any permutation of the roots is realized by an F-automorphism of K. On the other hand, the only subgroup of S_3 of order 3 is the alternating group A_3, a cyclic group. So if $[K : F] = 3$, then $G = A_3$. In this case the cyclic permutations and the identity are the only ones which extend to F-automorphisms. Thus the roots of an irreducible cubic polynomial may have either dihedral or cyclic symmetry. But these

symmetries are algebraic; they will not be symmetries of K when this field is viewed as a set of points in the complex plane.

Let us determine the intermediate fields in the case that the degree $[K : F]$ is 6. (There are no intermediate fields properly between F and K when $[K : F] = 3$.) The symmetric group S_3 has three conjugate subgroups of order 2 and one subgroup, A_3, of order 3. There are three obvious intermediate fields: $F(\alpha_1), F(\alpha_2), F(\alpha_3)$. They are isomorphic but not equal subfields of K, and they correspond to the three subgroups of order 2. But the intermediate field which corresponds to the subgroup A_3 is not obvious. Let us denote this mystery field by L. According to the Main Theorem, $G(K/L) = A_3$. Hence $[K : L] = 3$ and $[L : F] = 2$. So L is a quadratic extension of F, which can be obtained by adjoining a square root. The Main Theorem has told us an interesting fact: K contains the square root δ of an element of F. And since there is only one intermediate extension of degree 2, this square root is essentially unique. The Main Theorem also tells us that L is the fixed field of the subgroup A_3. So an even permutation of the roots leaves δ fixed, while an odd permutation does not. The required element is

$$\delta = (\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3)(\alpha_2 - \alpha_3). \tag{2.13}$$

A permutation of the roots multiplies δ by the sign of the permutation. Hence δ is not fixed by all elements of $G(K/F) = S_3$, so $\delta \notin F$. But δ^2 is fixed by every permutation. Corollary (1.9) tells us that $\delta^2 \in F$.

For any cubic polynomial $f(x) = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3)$, the element

$$D = (\alpha_1 - \alpha_2)^2(\alpha_1 - \alpha_3)^2(\alpha_2 - \alpha_3)^2 \tag{2.14}$$

is called the *discriminant* of the polynomial. It is an element of the field F which is zero if and only if two roots of $f(x)$ are equal. So it is analogous to the discriminant of the quadratic polynomial $x^2 + bx + c = (x - \alpha_1)(x - \alpha_2)$, which is $b^2 - 4c = (\alpha_1 - \alpha_2)^2$. If the cubic f is irreducible, then its roots are distinct, hence $D \neq 0$.

The fact that the discriminant of the cubic polynomial is an element of F follows from Corollary (1.9), but it is not trivial. We will prove it abstractly in the next section, but it can also be checked by direct calculation. Using formulas (2.8), we can compute the discriminant in terms of the coefficients p, q. It is

$$D = -4p^3 - 27q^2. \tag{2.15}$$

(2.16) **Proposition.** The discriminant of an irreducible cubic polynomial $f(x) \in F[x]$ is a square in F if and only if the degree of the splitting field is 3.

If we choose a polynomial with integer coefficients at random, the chances are good that its discriminant will not be a square in $\mathbb{Q}$. For example, the discriminant of $x^3 + 3x + 1$ is -135. On the other hand, the discriminant of $x^3 - 3x + 1$ is 81, a square. This agrees with the fact that $[K : F] = 3$ [see (2.12)].

Proof of the Proposition. If D is not a square, then $\delta \notin F$, and therefore $[F(\delta) : F] = 2$. Since $\delta \in K$, $[K : F]$ is divisible by 2, hence by (2.11), $[K : F] =$

6. On the other hand, if $\delta \in F$, then every element of the Galois group $G = G(K/F)$ fixes δ. Since odd permutations change the sign of δ, they are not in G, and hence $G \neq S_3$. Therefore $[K : F] = 3$. □

How could such a proposition be true? There must be a formula which expresses the second root α_2 in terms of the elements α_1, δ, and the coefficients p, q. This formula exists, and it is instructive to compute it explicitly.

3. SYMMETRIC FUNCTIONS

Galois theory is concerned with the problem of determining those permutations of the roots of a polynomial which extend to field automorphisms. In this section we examine a simple situation in which every permutation extends, namely when the roots are independent variables.

Let R be any ring, and consider the polynomial ring $R[u_1, \ldots, u_n]$ in n variables u_i. A permutation σ of $\{\mathbf{1}, \ldots, \mathbf{n}\}$ can be made to operate on polynomials, by permuting the variables. We must decide here how we want permutations to operate. Let us keep automorphisms on the left. Then σ operates by the inverse permutation on the indices:

$$f = f(u_1, \ldots, u_n) \overset{\sigma}{\rightsquigarrow} f(u_{1\sigma^{-1}}, \ldots, u_{n\sigma^{-1}}) = \sigma f. \tag{3.1}$$

This is clearly an automorphism of $R[u]$. Since it acts as the identity on R, σ is called an *R-automorphism*. So the symmetric group S_n operates by R-automorphisms on the polynomial ring $R[u]$. A polynomial is called *symmetric* if it is left fixed by all permutations.

It is easy to describe the symmetric polynomials. In order for g to be symmetric, two monomials in $\{u_1, \ldots, u_n\}$ which differ by a permutation of the indices, such as $u_1^2u_2$ and $u_2^2u_3$, must have the same coefficients in g. A symmetric polynomial which involves a given monomial must include the whole orbit. Thus

$$g(u) = (u_1^3+u_2^3+u_3^3) + 5(u_1^2u_2+u_1^2u_3+u_2^2u_3+u_2^2u_1+u_3^2u_2+u_3^2u_1) - u_1u_2u_3$$

is a symmetric polynomial of degree 3 in three variables.

There are n special symmetric polynomials with integer coefficients, called the *elementary symmetric functions* s_i:

$$\begin{aligned} s_1 &= u_1 + u_2 + \cdots + u_n \\ s_2 &= u_1u_2 + u_1u_3 + \cdots + u_{n-1}u_n = \sum_{i<j} u_iu_j \\ s_3 &= \sum_{i<j<k} u_iu_ju_k \\ &\vdots \\ s_n &= u_1u_2\cdots u_n. \end{aligned} \tag{3.2}$$

They are the coefficients of the polynomial $(x - u_1)(x - u_2)\cdots(x - u_n)$ when it is expanded as a polynomial in x:

(3.3) $p(x) = (x - u_1)(x - u_2)\cdots(x - u_n) = x^n - s_1x^{n-1} + s_2x^{n-2} - \cdots \pm s_n.$

We have reversed the order of the indices and alternated the sign here. The coefficients s_i are symmetric because $p(x)$ is symmetric with respect to permutation of the indices.

The main theorem on symmetric functions asserts that the elementary symmetric functions generate the ring of all symmetric polynomials:

(3.4) **Theorem.** Every symmetric polynomial $g(u_1,\dots,u_n) \in R[u]$ can be written in a unique way as a polynomial in the elementary symmetric functions $s_1,\dots,s_n$. In other words, let $z_1,\dots,z_n$ be variables. For each symmetric polynomial $g(u)$, there is a unique polynomial $\varphi(z_1,\dots,z_n) \in R[z_1,\dots,z_n]$ such that

$$g(u_1,\dots,u_n) = \varphi(s_1,\dots,s_n).$$

The proof of this theorem is at the end of the section.

For example,

$$u_1^2 + \cdots + u_n^2 = s_1^2 - 2s_2. \tag{3.5}$$

The *discriminant* of the polynomial $p(x)$ (3.3), defined to be

$$\begin{aligned} D &= (u_1 - u_2)^2(u_1 - u_3)^2\cdots(u_{n-1} - u_n)^2 \\ &= \prod_{i<j}(u_i - u_j)^2 = \pm\prod_{i\neq j}(u_i - u_j), \end{aligned} \tag{3.6}$$

is perhaps the most important symmetric polynomial. Both of the last two expressions for the discriminant are convenient at times, so it is unfortunate that they may differ by a sign. To go from the second expression for D to the last one requires $\frac{1}{2}n(n-1)$ sign changes, so the correct sign to replace the symbol $\pm$ is

$$(-1)^{n(n-1)/2}. \tag{3.7}$$

It is clear that D is a symmetric polynomial with integer coefficients. So Theorem (3.4) tells us that it can be written as an integer polynomial in the elementary symmetric functions. In other words, there exists a polynomial

$$\Delta(z_1,\dots,z_n) \in \mathbb{Z}[z_1,\dots,z_n] \tag{3.8}$$

so that $D = \Delta(s_1,\dots,s_n)$. Unfortunately, this expression for D in terms of the elementary symmetric functions is very complicated. I don't know what it is for $n > 3$.

We can compute the discriminant for $n = 2$ easily:

$$(u_1 - u_2)^2 = s_1^2 - 4s_2. \tag{3.9}$$

This is the familiar formula for the discriminant of the quadratic polynomial $p(x) = x^2 - s_1x + s_2$. When $n = 3$, the expression for the discriminant is already too complicated to remember:

(3.10)
$$(u_1 - u_2)^2(u_1 - u_3)^2(u_2 - u_3)^2 = s_1^2s_2^2 - 4s_2^3 - 4s_1^3s_3 - 27s_3^2 + 18s_1s_2s_3.$$

It is important to note that such an expression is an *identity* in $\mathbb{Z}[u_1,\dots,u_n]$. It remains true when substitutions are made for the variables u_i. If we are given particular elements $\{\alpha_1,\dots,\alpha_n\}$ in a ring R, we can expand the polynomial obtained by substituting α_i for u_i in $p(x)$:

$$(x - \alpha_1)(x - \alpha_2)\cdots(x - \alpha_n) = x^n - b_1x^{n-1} + b_2x^{n-2} - \cdots \pm b_n.$$

The indices and the signs have been adjusted to agree with (3.3). Then

$$b_i = s_i(\alpha_1,\dots,\alpha_n),$$

and

$$\prod_{i<j}(\alpha_i - \alpha_j)^2 = \Delta(b_1,\dots,b_n).$$

This follows by substitution of α_i for u_i.

It is also important that the expression of a symmetric polynomial in terms of the elementary symmetric functions is unique:

(3.11) **Corollary.** There are no polynomial relations among the elementary symmetric functions $s_1,\dots,s_n$. Equivalently, the subring $R[s_1,\dots,s_n]$ of $R[u]$ generated by $\{s_i\}$ is isomorphic to the polynomial ring $R[z_1,\dots,z_n]$ in n variables.

This is a restatement of the uniqueness in Theorem (3.4). □

The corollary can be used in the following way: Let

(3.12)
$$f(x) = x^n - a_1x^{n-1} + a_2x^{n-2} - \cdots \pm a_n$$

be a polynomial with coefficients in a ring R. We define the *discriminant* of $f(x)$ to be the element $\Delta(a_1,\dots,a_n)$ of R, where $\Delta(z_1,\dots,z_n)$ is the polynomial (3.8). Since this polynomial is unique, the discriminant is defined, whether the polynomial is a product of linear factors in $R[x]$ or not.

For example, let $n = 3$. Then formula (3.10) shows that

(3.13)
$$\Delta(0, p, -q) = -4p^3 - 27q^2,$$

which agrees with the formula (2.15) for the discriminant of the cubic polynomial $x^3 + px + q$.

We can use undetermined coefficients to compute the expression of a symmetric polynomial in terms of the elementary symmetric functions. To apply this

method, we notice that the elementary symmetric function s_i has degree i in the variables u. That is why we chose the index i for it. So we assign the *weight* i to the variable z_i, and we define the *weighted degree* of a monomial $z_1^{e_1}z_2^{e_2}\cdots z_n^{e_n}$ to be

$$e_1 + 2e_2 + \cdots + ne_n. \tag{3.14}$$

Substitution of s_i for z_i into a polynomial of weighted degree d in z yields a polynomial of (ordinary) degree d in $u_1,\dots,u_n$.

For example, to compute the discriminant of a cubic polynomial in terms of the elementary symmetric functions, we notice that its degree in u is 6. There are seven monomials in z_1, z_2, z_3 of weighted degree 6:

$$z_1^6,\ z_1^4z_2,\ z_1^3z_3,\ z_1^2z_2^2,\ z_1z_2z_3,\ z_2^3,\ z_3^2. \tag{3.15}$$

So D is a linear combination of these monomials. To compute its coefficients, we evaluate D on some special polynomials: Setting $f(x) = x^2(x-1)$, we get $D = 0$, $s_1 = 1$, and $s_2 = s_3 = 0$. Since the only one of the monomials (3.15) which does not involve z_2 or z_3 is z_1^6, the coefficient of z_1^6 in the discriminant is zero. The coefficients of z_2^3 and z_3^2 can be computed using the special polynomials $x^3 - x$ and $x^3 - 1$, for example.

Proof of Theorem (3.4). Let's warm up by working out the case of the symmetric polynomial

$$f(x) = u_1^2u_2 + u_1^2u_3 + u_2^2u_1 + u_2^2u_3 + u_3^2u_1 + u_3^2u_2$$

as an example. To analyze it, our first step is to set $u_3 = 0$. We obtain a symmetric polynomial $f^0 = u_1^2u_2 + u_2^2u_1$ in the remaining variables u_1, u_2. Let us denote the elementary symmetric functions in u_1, u_2 by $s_1^0 = u_1 + u_2$ and $s_2^0 = u_1u_2$. We notice that $f^0 = s_1^0s_2^0$.

The second step is to compare f with the polynomial s_1s_2 in three variables. We compute the polynomial $f - s_1s_2$, where $s_1 = u_1 + u_2 + u_3$ and $s_2 = u_1u_2 + u_1u_3 + u_2u_3$, finding that

$$f - s_1s_2 = -3u_1u_2u_3.$$

We recognize this polynomial as $-3s_3$. So $f = s_1s_2 - 3s_3$.

The general case is similar. There is nothing to show when $n = 1$, because $u_1 = s_1$ in that case. Proceeding by induction, we assume the theorem proved for $n-1$ variables. Given a symmetric polynomial f in $u_1,\dots,u_n$, we consider the polynomial f^0 obtained by substituting zero for the last variable: $f^0(u_1,\dots,u_{n-1}) = f(u_1,\dots,u_{n-1},0)$. We note that f^0 is a symmetric polynomial in $u_1,\dots,u_{n-1}$. By the induction hypothesis, f^0 may be expressed as a polynomial in the elementary symmetric functions in $\{u_1,\dots,u_{n-1}\}$, which we denote by

$$s_1^0 = u_1 + \cdots + u_{n-1}, \dots, s_{n-1}^0 = u_1\cdots u_{n-1}.$$

So we can write $f^0 = g(s_1^0,\dots,s_{n-1}^0)$. Moreover, it follows from the defintion of the polynomials s_i that

$$s_i^0 = s_i(u_1,\dots,u_{n-1},0), \quad \text{if } i = 1,\dots, n-1.$$

Consider the polynomial

$$p(u_1,\ldots,u_n) = f(u_1,\ldots,u_n) - g(s_1,\ldots,s_{n-1}),$$

as a polynomial in $u_1,\ldots,u_n$. Being a difference of symmetric polynomials, this polynomial is symmetric. Also, it has the property that $p(u_1,\ldots,u_{n-1},0) = 0$. Therefore every monomial occurring in p is divisible by u_n. By symmetry, p is divisible by u_i for every i, and hence it is divisible by s_n. So

$$f(u_1,\ldots,u_n) = g(s_1,\ldots,s_{n-1}) + s_n h(u_1,\ldots,u_n), \tag{3.16}$$

for some symmetric polynomial h. We now work on $h(u_1,\ldots,u_n)$. By induction on the degree, h is a polynomial in the symmetric functions, and hence so is f.

It remains to prove the uniqueness of $\varphi(s_1,\ldots,s_n)$. The uniqueness means that there is only one polynomial $\varphi(z_1,\ldots,z_n)$ in the variables z_i, such that $\varphi(s_1,\ldots,s_n) = f(u_1,\ldots,u_n)$, as polynomials in $u_1,\ldots,u_n$. In other words, the kernel of the substitution map

$$\sigma: R[z] \longrightarrow R[u]$$

sending $z_i \rightsquigarrow s_i$ is zero. To show this, suppose $\varphi(s_1,\ldots,s_n) = 0$ for some $\varphi \in R[z]$. Setting $u_n = 0$ in this expression we still get zero: $\varphi(s_1{}^0,\ldots,s_{n-1}{}^0,0) = 0$. By induction on n, this implies that $\varphi(z_1,\ldots,z_{n-1},0) = 0$. Therefore z_n divides $\varphi(z)$, and we may write $\varphi(z) = z_n\psi(z)$. Then $0 = \varphi(s) = s_n\psi(s) = u_1\cdots u_n\psi(s)$. Since the product $u_1\cdots u_n$ is not a zero-divisor in the polynomial ring $R[u]$, $\psi(s) = 0$. The polynomial $\psi(z)$ has lower total degree in z than $\varphi(z)$, so we may apply induction on the degree to conclude that $\psi = 0$. Hence $\varphi = 0$ too. □

Now suppose that $R = F$ is a field. Then we may also consider the field of rational functions in the variables u_i, that is, the field of fractions of $F[u_1,\ldots,u_n]$. The symmetric group also acts on this field, and the corresponding assertion is true:

(3.17) **Theorem.** Every symmetric rational function is a rational function in $s_1,\ldots,s_n$.

Proof. Let $r(u) = f(u)/g(u)$ be a symmetric rational function, where $f,g \in F[u]$. We can build a symmetric function from g by multiplying all the σg together:

$$G = \prod_{\sigma\in S_n} \sigma g$$

is a symmetric polynomial. Then $G(u)r(u)$ is a symmetric rational function, and it is also a polynomial in $\{u_1,\ldots,u_r\}$—a symmetric polynomial. By Theorem (3.4), $G(u)$ and $G(u)r(u)$ are polynomials in the elementary symmetric functions $\{s_i\}$. Thus $r(u)$ is a rational function in $\{s_i\}$. □

The pair of fields

$$F(s) = F(s_1,\ldots,s_n) \subset F(u_1,\ldots,u_n) = F(u) \tag{3.18}$$

is an example of a Galois extension. This follows from Theorem (1.11), because $F(u)$ is a splitting field of the polynomial $p(x)$ (3.3) and because the roots $u_1,\dots,u_n$ are distinct. By Proposition (1.14), the Galois group $G = G(F(u)/F(s))$ operates faithfully on the roots. On the other hand, G contains the full symmetric group, by construction. Therefore $G = S_n$. As a corollary, we find that $[F(u):F(s)] = n!$. Needless to say, this can be proved directly.

4. PRIMITIVE ELEMENTS

At the end of the first section, we saw that generically chosen elements of a biquadratic extension K/F generate K. It is possible to derive a general statement of this type as a corollary of the Main Theorem of Galois theory. But we are going to prove it directly instead, and then use this fact in the proof of the Main Theorem.

(4.1) **Theorem.** *Existence of a primitive element:* Let K be a finite extension of a field F of characteristic zero. There is an element $\gamma \in K$ such that $K = F(\gamma)$.

An element γ which generates a field extension K/F is called a *primitive element* for K over F. So the theorem can be restated by saying that every finite extension K of a field F has a primitive element. We have restated our general hypothesis that F has characteristic zero here because this theorem is not true for fields of characteristic p.

Proof of Theorem (4.1). We use induction on the number of generators of K. Say that $K = F(\alpha_1,\dots,\alpha_n)$. If $n = 1$, there is nothing to prove. For $n > 1$, the induction principle allows us to assume the theorem true for the intermediate field $K_1 = F(\alpha_1,\dots,\alpha_{n-1})$. So we may assume that K_1 is generated by a single element β. Then $K = K_1(\alpha_n) = F(\beta,\alpha_n)$. We have to show that this field has a primitive element. We are thereby reduced to the case that $n = 2$, so that K is generated by two elements α,β.

Let $f(x), g(x)$ be the irreducible polynomials for α,β over F, and let K' be an extension of K in which f and g split completely [Chapter 13 (5.3)]. Call their roots $\alpha = \alpha_1,\dots,\alpha_m$ and $\beta = \beta_1,\dots,\beta_n$. By Chapter 13 (5.8), the elements α_i are distinct.

We are going to show that for most choices of $c \in F$, the linear combination $\gamma = \beta + c\alpha$ generates K. Let us denote the field $F(\gamma)$ by L. It suffices to show that $\alpha \in L$, because if so, then $\beta = \gamma - c\alpha$ will be in L too, and this will imply that $L = K$. The way we show that α is in L is indirect: We determine its irreducible polynomial over L. As we know, this is the monic polynomial of least degree in $L[x]$ which has α as a root.

To begin with, α is a root of $f(x)$. The trick is to use the polynomial $g(x)$ to cook up a second polynomial with the root α, namely $h(x) = g(\gamma - cx)$. Notice that $h(x)$ has coefficients in L and that $h(\alpha) = 0$. If we show that the greatest com-

mon divisor of f and h in $L[x]$ is $x - \alpha$, then it will follow that $-\alpha$, being one of the coefficients of $x - \alpha$, is in L. Now the monic greatest common divisor of f and h is the same, whether computed in $L[x]$ or in $K'[x]$ [Chapter 13 (5.4)]. So we may make our computation in $K'[x]$. In that ring, f is a product of the linear factors $x - \alpha_i$, and it suffices to show that none of them divides h, that is, that none of the elements α_i, except for $\alpha = \alpha_1$ itself, is a root of $h(x)$. Having gotten this far, the rest is just a matter of computing the roots of h.

Since the roots of g are β_j, the roots of $h(x) = g(\gamma - cx)$ are obtained by solving the equations

$$\gamma - cx = \beta_j$$

for x. Since $\gamma = \beta + c\alpha$, the roots are $(\gamma - \beta_j)/c = (\beta - \beta_j)/c + \alpha$. We want these roots to be different from α_i, $i \neq 1$. This will be so provided that c does not take one of the finitely many values

$$-\frac{\beta_j - \beta}{\alpha_i - \alpha}, \tag{4.2}$$

with $i, j \neq 1, 1$. □

(4.3) **Example.** Consider the field $K = \mathbb{Q}[i, \sqrt[3]{2}]$. This field has degree 6 over $\mathbb{Q}$ [see Chapter 13 (3.5d)]. In the notation of the previous proof, we have $\beta_1 = i$, $\beta_2 = -i$, and $\alpha_1 = \sqrt[3]{2}$, $\alpha_2 = \zeta\sqrt[3]{2}$, $\alpha_3 = \zeta^2\sqrt[3]{2}$, where $\zeta = e^{2\pi i/3}$. Condition (4.2) becomes

$$\sqrt[3]{2}c \neq -\frac{\pm i - i}{\zeta^\nu - \zeta}, \quad \nu = 1, 2.$$

This condition holds for all $c \in \mathbb{Q}$ except $c = 0$. Therefore $\gamma = i + c\sqrt[3]{2}$ generates K over $\mathbb{Q}$ for all rational numbers $c \neq 0$. Of course, many other combinations of the two elements β, α will generate $F(\beta, \alpha)$. In this example, the product $i\sqrt[3]{2}$ also generates K. □

Theorem (4.1) is important for two reasons. First, explicit computation in an extension of the form $F(\gamma)$ is easy if the irreducible equation for γ over F is known. Second, since finite extensions have the form $F(\gamma)$, we can derive their properties from facts about algebraic elements. It is this aspect which is most important for us.

The power of Theorem (4.1) is shown by applying it to the study of automorphisms of fields. Consider a finite group G of automorphisms of the field K, and denote its fixed field K^G by F.

(4.4) **Proposition.** Let G be a finite group of automorphisms of a field K, and let F be its fixed field. Let $\{\beta_1, \ldots, \beta_r\}$ be the orbit of an element $\beta = \beta_1 \in K$ under the action of G. Then β is algebraic over F, its degree over F is r, and its irreducible polynomial over F is $g(x) = (x - \beta_1) \cdots (x - \beta_r)$.

Note that the degree of β, being the order of an orbit, divides the order of the group.

Proof. Let $f(x)$ be the irreducible polynomial for β over F. Since $f(x)$ is fixed by G, each of the elements β_i is a root of f (1.14), and so g divides f. Also, g is fixed by all permutations of $\{\beta_1,\dots,\beta_r\}$, and hence by the operation of G, which permutes the orbit. Therefore $g(x) \in F[x]$. Since f is irreducible, $g = f$. □

This proposition provides a method for determining the irreducible polynomial for an element β of a Galois extension K over F. For example, let K be the biquadratic extension $\mathbb{Q}(i, \sqrt{2})$, and let $\beta = i + \sqrt{2}$. The Galois group of $K/\mathbb{Q}$ is the Klein four group, and the orbit of β consists of the four elements $\pm i \pm \sqrt{2}$. So the irreducible polynomial for β over $\mathbb{Q}$ is

$$(x - i - \sqrt{2})(x - i + \sqrt{2})(x + i - \sqrt{2})(x + i + \sqrt{2})$$
$$= (x^2 - 2ix - 3)(x^2 + 2ix - 3) = x^4 - 2x^2 + 9.$$

We can also determine this polynomial by computing powers of β and finding the linear relation of smallest degree between them (see Chapter 13, Section 3). However, the method given here is preferable because it always produces an irreducible polynomial.

(4.5) **Corollary.** Let K/F be a Galois extension, and let $g(x)$ be an irreducible polynomial in $F[x]$. If g has one root in K, then it factors into linear factors in $K[x]$.

Proof. According to Corollary (1.9), F is the fixed field of the Galois group $G = G(K/F)$. Let β be a root of $g(x)$ in K. By Proposition (4.4), the irreducible polynomial for β over F is $(x - \beta_1)\cdots(x - \beta_r)$, where $\{\beta_1,\dots,\beta_r\}$ is the G-orbit of β. Since $g(x)$ is the irreducible polynomial for β, it is equal to this product, so it factors into linear factors in K, as asserted. □

The corollary tells us in particular that every Galois extension is a splitting field, which is part of Theorem (1.11). For, take any generators $\alpha, \beta, \dots$ for K over F, and let $f(x)$ be the product of their irreducible polynomials. Then f splits completely in K, and hence K is a splitting field for f.

(4.6) **Theorem.** Let G be a group of order n of automorphisms of a field K, and let F be its fixed field. Then $[K : F] = n$.

Proof. Proposition (4.4) shows that every element β of K is algebraic over F and that its degree divides $n = |G|$. The theorem of the primitive element implies that the degree of the whole field extension K/F is bounded by n too. To see this, we form a chain of extension fields as follows: We choose an element $\alpha_1 \in K$ which is not in F, and we set $F_1 = F(\alpha_1)$. Then $[F_1 : F] \leq n$. If $F_1 \neq K$, we choose an element $\alpha_2 \in K$ which is not in F_1, and we set $F_2 = F(\alpha_1, \alpha_2)$. By the theorem of the primitive element, F_2 is generated by a single element γ, and by Corollary (3.6) of

Chapter 13, the degree of γ over F is bounded by n. So $[F_2 : F] \leq n$. Continuing in this way, we obtain a chain $F < F_1 < F_2 \ldots$ in which $[F_i : F] \leq n$ for all i. This chain must be finite. So $F_i = K$ for some i, and $[K : F] \leq n$.

Applying Theorem (4.1) once more, we conclude that K has a primitive element: $K = F(\beta)$. Any element of G which fixes β acts as the identity on $K = F(\beta)$. Since we are assuming that G is a group of automorphisms of K, the identity is the only such element. Therefore the stabilizer of β is $\{1\}$, and the orbit has order n. By Proposition (4.4), β has degree n over F, and $[K : F] = n$. □

Using the theorem we have just proved, we can derive the first theorem, Theorem (1.6), which was stated in Section 1. That theorem says that for any finite extension K/F, the order of its Galois group divides its degree. To prove this, we set $G = G(K/F)$. Then G operates on K, so by Theorem (4.6), $|G| = [K : K^G]$. And since $F \subset K^G \subset K$, $[K : K^G]$ divides $[K : F]$. □

Theorem (4.6) also provides us with a converse to Corollary (1.9):

(4.7) **Corollary.** Let G be a finite group of automorphisms of a field K, and let F be its fixed field. Then K is a Galois extension of F, and its Galois group is G.

Proof. By definition of the fixed field, the elements of G are F-automorphisms of K. Hence $G \subset G(K/F)$. Since $|G(K/F)| \leq [K : F]$ and $[K : F] = |G|$, it follows that $|G(K/F)| = [K : F]$ and that $G = G(K/F)$. □

We can get some interesting examples to illustrate Proposition (4.4) and Theorem (4.6) by considering automorphisms of the field $\mathbb{C}(y) = K$ of rational functions in y. For instance, let σ, τ be the automorphisms of K defined by $y \rightsquigarrow -y$ and $y \rightsquigarrow iy^{-1}$. The automorphisms $\{1, \sigma, \tau, \sigma\tau\}$ form a group G of order 4.

(4.8) **Proposition.** Let K and G be as above. The fixed field $F = K^G$ is the field $\mathbb{C}(w)$ of rational functions in $w = y^2 - y^{-2}$.

In other words, every rational function $f(y)$ which is fixed by σ can be expressed as a rational function in w.

Proof. First of all, G does fix $w = y^2 - y^{-2}$, so w is in the fixed field. Therefore the fixed field F contains the field $\mathbb{C}(w)$. Next, we compute the irreducible polynomial for y over F. The orbit of y is $\{y, iy^{-1}, -y, -iy^{-1}\}$, so Proposition (4.4) tells us that the irreducible equation for y is $(x - y)(x - iy^{-1})(x + y)(x + iy^{-1}) = x^4 - wx^2 - 1$. This polynomial has coefficients in $\mathbb{C}(w)$, so y has degree 4 over that field. It follows that $[K : \mathbb{C}(w)] = 4$. On the other hand, $\mathbb{C}(w) \subset F \subset K$, and since $|G| = 4$, Theorem (4.6) tells us that $[K : F] = 4$. Counting degrees shows that $\mathbb{C}(w) = F$. □

A famous theorem called *Lüroth's theorem* asserts that any subfield of the field $\mathbb{C}(y)$ which properly contains the complex numbers is the field of rational functions in some rational function w of y.

5. PROOF OF THE MAIN THEOREM

Let $f(x)$ be a monic polynomial of degree n with coefficients in a field F. We recall that a splitting field of $f(x) \in F[x]$ is a field of the form $K = F(\alpha_1, \ldots, \alpha_n)$, such that $f(x) = (x - \alpha_1) \cdots (x - \alpha_n)$ in $K[x]$. The existence of a splitting field was proved in Chapter 13 (5.3). We now want to show that any two splitting fields of a given polynomial $f(x)$ are isomorphic. This follows from the fact that a field extension of the form $F(\alpha)$ is determined by the irreducible polynomial for α over F, and from some "bookkeeping." The bookkeeping required for the proof is notationally a little confusing, but not difficult.

Any isomorphism $\varphi: F \longrightarrow \tilde{F}$ of fields extends to an isomorphism $F[x] \longrightarrow \tilde{F}[x]$ between the polynomial rings by

$$a_n x^n + \cdots + a_0 \rightsquigarrow \tilde{a}_n x^n + \cdots + \tilde{a}_0,$$

where $\tilde{a}_i = \varphi(a_i)$. Let us denote the image of $f(x)$ by $\tilde{f}(x)$. Since φ is an isomorphism, $\tilde{f}(x)$ will be an irreducible polynomial if and only if $f(x)$ is irreducible.

The following lemma generalizes Chapter 13 (2.9).

(5.1) **Lemma.** With the above notation, let $f(x)$ be an irreducible polynomial in $F[x]$. Let α be a root of $f(x)$ in an extension field K of F, and let $\tilde{\alpha}$ be a root of $\tilde{f}(x)$ in an extension $\tilde{K}$ of $\tilde{F}$. There is a unique isomorphism

$$\varphi_1: F(\alpha) \longrightarrow \tilde{F}(\tilde{\alpha})$$

which restricts to φ on the subfield F, and which sends α to $\tilde{\alpha}$.

Proof. We know that $F(\alpha)$ is isomorphic to the quotient $F[x]/(f)$, and similarly $\tilde{F}(\tilde{\alpha})$ is isomorphic to $\tilde{F}[x]/(\tilde{f})$. The rings $F[x]$ and $\tilde{F}[x]$ are isomorphic, as we just saw, and since f and $\tilde{f}$ correspond under this isomorphism, so do the ideals (f) and $(\tilde{f})$ which they generate. Therefore the residue rings $F[x]/(f)$ and $\tilde{F}[x]/(\tilde{f})$ are also isomorphic. Combining these isomorphisms yields the required isomorphism φ_1. This extension of φ is unique because α generates $F(\alpha)$ over F. □

(5.2) **Proposition.** Let $\varphi: F \longrightarrow \tilde{F}$ be an isomorphism of fields. Let $f(x)$ be a nonconstant polynomial in $F[x]$, and let $\tilde{f}(x)$ be the corresponding polynomial in $\tilde{F}[x]$. Let K and $\tilde{K}$ be splitting fields for $f(x)$ and $\tilde{f}(x)$. There is an isomorphism $\psi: K \longrightarrow \tilde{K}$ which restricts to φ on the subfield F of K.

If we let $F = \tilde{F}$ and $\varphi =$ identity, we obtain the following corollary:

(5.3) **Corollary.** Any two splitting fields of $f(x) \in F[x]$ over F are isomorphic. □

The corollary is the result we are really after. The auxiliary isomorphism φ is introduced into the proposition to make the induction step of the proof work.

Proof of Proposition (5.2). If $f(x)$ factors into linear factors over F, then $\tilde{f}(x)$ also factors into linear factors. In this case $K = F$ and $\tilde{K} = \tilde{F}$, so $\varphi = \psi$. Assume that f does not split completely. Choose an irreducible factor $g(x)$ of $f(x)$ of degree >1. The corresponding polynomial $\tilde{g}(x)$ will be an irreducible factor of $\tilde{f}(x)$. Let α be a root of g in K and write $F_1 = F(\alpha)$. Make a similar choice of $\tilde{\alpha}$ and $\tilde{F}_1 = \tilde{F}(\tilde{\alpha})$ in $\tilde{K}$. Then by Lemma (5.1), we can extend φ to an isomorphism $\varphi_1: F_1 \longrightarrow \tilde{F}_1$ which sends $\alpha \rightsquigarrow \tilde{\alpha}$. Being a splitting field for f over F, K is also a splitting field of f over the larger field F_1, and similarly $\tilde{K}$ is a splitting field for $\tilde{f}$ over $\tilde{F}_1$. Therefore we may replace $F, \tilde{F}, \varphi$ by $F_1, \tilde{F}_1, \varphi_1$ and proceed by induction on the degree of K over F. □

We are now in a position to prove the second of the theorems, Theorem (1.11), which was announced in Section 1. One part of this theorem was proved in the last section, using Corollary (4.5). For convenience, we restate the other part here.

Theorem. Let K be the splitting field of a polynomial $f(x) \in F[x]$. Then K is a Galois extension of F; that is, $|G(K/F)| = [K : F]$.

We will prove the theorem by going back over the proof of Proposition (5.2), keeping careful track of the number of choices.

(5.4) **Lemma.** With the notation of (5.2), the number of isomorphisms $\psi: K \longrightarrow \tilde{K}$ extending φ is equal to the degree $[K : F]$.

The theorem follows from this lemma if we set $\tilde{F} = F$, $\tilde{K} = K$, and $\varphi =$ identity. □

Proof of Lemma (5.4). We proceed as in the proof of Proposition (5.2), choosing an irreducible factor $g(x)$ of $f(x)$ and one of the roots α of $g(x)$ in K. Let $F_1 = F(\alpha)$. Any isomorphism $\psi: K \longrightarrow \tilde{K}$ extending φ will send F_1 to some subfield $\tilde{F}_1$ of $\tilde{K}$. This field $\tilde{K}$ will have the form $\tilde{F}(\tilde{\alpha})$, where $\tilde{\alpha} = \psi(\alpha)$ is a root of $\tilde{g}(x)$ in $\tilde{K}$.

Conversely, to extend φ to ψ, we may start by choosing any root $\tilde{\alpha}$ of $\tilde{g}(x)$ in $\tilde{K}$. We then extend φ to a map $\varphi_1: F_1 \longrightarrow \tilde{F}_1 = \tilde{F}(\tilde{\alpha})$ by setting $\varphi_1(\alpha) = \tilde{\alpha}$. We use induction on $[K : F]$. Since $[K : F_1] < [K : F]$, the induction hypothesis tells us that for this particular choice of φ_1, there are $[K : F_1]$ extensions of φ_1 to an isomorphism $\psi: K \longrightarrow \tilde{K}$. On the other hand, $\tilde{g}$ has distinct roots in $\tilde{K}$ because g and $\tilde{g}$ are irreducible [Chapter 13 (5.8)]. So the number of choices for $\tilde{\alpha}$ is the degree of g, which is $[F_1 : F]$. There are $[F_1 : F]$ choices for the isomorphism φ_1. This gives us a total of $[K : F_1][F_1 : F] = [K : F]$ extensions of φ to $\psi: K \longrightarrow \tilde{K}$. □

Since any two splitting fields K of a polynomial $f(x) \in F[x]$ are isomorphic, the Galois group $G(K/F)$ depends, up to isomorphism, only on f. It is often referred to as the *Galois group of the polynomial* over F.

The following corollary collects together several criteria for an extension to be Galois. Most of them have already been proved, and we leave the remaining proofs as exercises.

(5.5) **Corollary.** Let K/F be a finite field extension. The following are equivalent:

(i) K is a Galois extension of F;
(ii) K is the splitting field of an irreducible polynomial $f(x) \in F[x]$;
(ii′) K is the splitting field of a polynomial $f(x) \in F[x]$;
(iii) F is the fixed field for the action of the Galois group $G(K/F)$ on K;
(iii′) F is the fixed field for an action of a finite group of automorphisms of K. □

We now have enough information to prove the Main Theorem of Galois theory, which relates intermediate fields to subgroups of the Galois group.

Proof of Theorem (1.15). Let K/F be a Galois extension. We have to show that the maps

$$L \rightsquigarrow G(K/L) \quad \text{and} \quad H \rightsquigarrow K^H$$

are inverse functions between the set of intermediate fields and the set of subgroups of $G = G(K/F)$. To do so, we verify that the composition of these two maps in either direction is the identity.

Let L be an intermediate field. The corresponding subgroup of G is $H = G(K/L)$. By definition, H acts trivially on L, so $L \subset K^H$. On the other hand, K is a Galois extension of L by (1.13); hence $[K:L] = |H|$. By Theorem (4.6), $|H| = [K:K^H]$, so $L = K^H$.

In the other direction, suppose that we start with a subgroup $H \subset G$, and let $L = K^H$. Then $H \subset G(K/L)$. But $|H| = [K:K^H] = [K:L] = |G(K/L)|$. Therefore $H = G(K/L)$. This shows that the two maps are inverses, as required. Since K is a Galois extension of $L = K^H$, $[K:L] = |H|$, and $[L:F] = [G:H]$. □

The correspondence given by the Main Theorem has some surrounding details which we will now discuss. First of all, the correspondence between fields and subgroups is *order reversing*, that is, if L, L' are two intermediate fields and if $H = G(K/L)$, $H' = G(K/L')$ are the corresponding subgroups, then $L \subset L'$ if and only if $H \supset H'$. This is clear from the definitions of the maps and is consistent with the relations (1.16).

To complete the picture, we will show that the immediate fields L which are *Galois extensions* of F correspond to the *normal subgroups* of G. Let L be an intermediate field. An F-automorphism σ of K will carry L to some intermediate field σL which may or may not be the same as L. We call σL a *conjugate subfield.*

(5.6) **Theorem.** Let K/F be a Galois extension, and let L be an intermediate field. Let $H = G(K/L)$ be the corresponding subgroup of $G = G(K/F)$.

(a) Let σ be an element of G. The subgroup of G which corresponds to the conjugate subfield σL is the conjugate subgroup $\sigma H \sigma^{-1}$. In other words, $G(K/\sigma L) = \sigma H \sigma^{-1}$.

(b) L is a Galois extension of F if and only if H is a normal subgroup of G. When this is so, then $G(L/F)$ is isomorphic to the quotient group G/H:

(5.7) **Diagram.**

$G = G(K/F)$ *operates on* K *fixing* F

K — L — F

$H = G(K/L)$ *operates on* K, *fixing* L (between K and L)

If H *is normal, then* $G/H = G(L/F)$ *operates here* (between L and F)

(5.8) **Example.** *In the case of the cubic equation* (2.1) *whose splitting field has de-*gree 6, the only intermediate extension which is Galois, other than F and K, is $F(\delta)$, which corresponds to the alternating group $H = A_3 \subset S_3$. The Galois group $G(F(\delta)/F)$ is cyclic of order 2, as is the quotient group S_3/A_3. The three fields $F(\alpha_i)$ are conjugate. This agrees with the fact that the three subgroups of S_3 of order 2 are conjugate.

Proof of Theorem (5.7). (a) Let $\sigma L = L'$. If τ is an element of $H = G(K/L)$, then $\sigma\tau\sigma^{-1}$ is in $H' = G(K/L')$. To check this, we must show that $\sigma\tau\sigma^{-1}$ fixes any element $\alpha' \in L'$. By definition of σL, $\alpha' = \sigma(\alpha)$ for some $\alpha \in L$. Then $\sigma\tau\sigma^{-1}(\alpha') = \sigma\tau(\alpha) = \sigma(\alpha) = \alpha'$, as required. It follows that $H' \supset \sigma H \sigma^{-1}$ and by symmetry, or by counting elements, that $H' = \sigma H \sigma^{-1}$. The fact which we have just checked is actually a general property of group actions on sets [Chapter 5 (6.4)].

(b) Now suppose that H is normal. Then $H = \sigma H \sigma^{-1}$ for all $\sigma \in G$; hence $G(K/L) = G(K/\sigma L)$. This implies that $L = \sigma L$ for all σ [see (1.9)]. Thus every F-automorphism of K carries L to itself and hence defines an F-automorphism of L by restriction. This restriction defines a homomorphism

(5.9) $$\pi: G \longrightarrow G(L/F).$$

Its kernel is the set of $\sigma \in G$ which induces the identity on L, which is H. Therefore G/H is isomorphic to a subgroup of $G(L/F)$. Counting degrees and orders, we find

$$[L : F] = |G/H| \leq |G(L/F)|.$$

It follows that L is a Galois extension and that $G/H \approx G(L/F)$.

Conversely, suppose that L/F is Galois. Then L is a splitting field of some polynomial $g(x) \in F[x]$; that is, $L = F(\beta_1, \ldots, \beta_k)$, where β_i are the roots of $g(x)$

in K. An F-automorphism σ of K permutes these roots and therefore carries L to itself: $L = \sigma L$. By (a), $H = \sigma H \sigma^{-1}$; thus H is a normal subgroup. □

6. QUARTIC EQUATIONS

Let K/F be a Galois extension. We have seen that if β is an element of K whose monic irreducible polynomial over F is $g(x)$, then g splits completely in K, and the G-orbit of β is the set of roots of g (4.4). So G operates *transitively* on the roots of an irreducible polynomial $g \in F[x]$, provided that this polynomial has at least one root in K. Combining this observation with Proposition (1.14), we find:

(6.1) **Proposition.** Let K/F be a splitting field of a polynomial $f(x) \in F[x]$. The Galois group G of K/F operates faithfully on the set $\{\alpha_1, \dots, \alpha_n\}$ of roots of f. Hence this operation represents G as a subgroup of the symmetric group S_n. The roots form a single orbit if and only if f is irreducible over F. □

When the Galois extension K is exhibited as the splitting field of a polynomial of degree n, it is customary to view the Galois group G as a subgroup of the symmetric group S_n. If the polynomial f is irreducible, then it is a *transitive* subgroup, which means that it acts transitively on the indices $\{\mathbf{1}, \dots, \mathbf{n}\}$. However, the same Galois extension K/F can be exhibited as a splitting field of many polynomials, so this representation of G as a subgroup of S_n is not unique.

For instance, let K/F be the splitting field of an irreducible cubic equation such that $[K : F] = 6$. Then the Galois group is represented as the whole symmetric group S_3. However, the theorem of the primitive element tells us that K can also be generated by a single element γ. Since $[K : F] = 6$, γ has degree 6 over F. This means that its orbit has order 6 and that its irreducible polynomial has degree 6. So if we think of K as the splitting field of this sextic polynomial, the Galois group is represented as a subgroup of S_6. This isn't a very economical way to represent S_3.

Let us suppose that our Galois extension K is the splitting field of a polynomial $f(x)$ and that its roots in K are $\alpha_1, \dots, \alpha_n$. Then, viewing G as a subgroup of S_n, we may pose the following two problems:

(6.2) (i) Given a subgroup $\mathcal{H}$ of S_n, decide if $G \subset \mathcal{H}$.
(ii) Determine G.

If we could solve (i) for every subgroup $\mathcal{H}$, then (ii) would also be solved.

Lagrange's approach to these problems is to look for functions of the roots which are *partially symmetric*. A partially symmetric polynomial is a polynomial $p(u_1, \dots, u_n)$ in the variables $\{u_1, \dots, u_n\}$ which is left fixed by the permutations in a given subgroup $\mathcal{H}$ of S_n but not by any other permutations. For example, we saw in (2.13) that

$$(u_1 - u_2)(u_1 - u_3)(u_2 - u_3)$$

is a partially symmetric function for the alternating group, when $n = 3$. There is no difficulty in generalizing this construction to arbitrary n by defining

(6.3) $$\delta(u) = (u_1 - u_2)(u_1 - u_3)\cdots(u_{n-1} - u_n) = \prod_{i<j}(u_i - u_j).$$

This element is a square root of the discriminant (3.6). The effect of a permutation of the indices is to multiply δ by the sign of the permutation. Having this partially symmetric function in hand, we substitute the roots $\alpha_1,\dots,\alpha_n$ of our polynomial into it, to obtain an element $\delta(\alpha) = \delta$ of K which is fixed by even permutations of the roots. We can decide whether or not δ is in F by determining whether or not the discriminant D is a square. This will provide information about the Galois group.

(6.4) **Proposition.** Let K/F be a Galois extension which is the splitting field of an irreducible polynomial $f(x) \in F[x]$ of degree n. Let $\alpha_1,\dots,\alpha_n$ be the roots of $f(x)$ in K, and let $\delta = \delta(\alpha)$. Then $\delta \neq 0$. Moreover:

(a) *$\delta \in F$ if and only if the Galois group G is a subgroup of the alternating group A_n.*

(b) In any case, the subgroup $G(K/F(\delta))$ of G is contained in the alternating group.

Proof. The case $\delta = 0$ occurs only if two of the roots are equal, and this can not happen if f is irreducible [Chapter 13 (5.8)]. Next, assume that δ is in F. Since odd permutations send $\delta \rightsquigarrow -\delta$ and since $\delta \neq 0$, odd permutations don't fix δ. On the other hand, the elements of F are fixed by every automorphism in G. It follows that G does not contain any odd permutations, hence that $G \subset A_n$. Conversely, if $\delta \notin F$, we use the fact that $K^G = F$. There must be an element of G which doesn't fix δ. This element will be an odd permutation, so $G \not\subset A_n$. This proves (a). Part (b) follows from (a) when we replace F by $F(\delta)$. □

We will now discuss quartic equations, beginning with an interesting special case which is controlled by the discriminant. We consider a complex number which is presented as a nested square root, say $\alpha = \sqrt{r+s\sqrt{t}}$, where r, s, t are in a field F. The numbers

(6.5) $$\sqrt{3+2\sqrt{2}},\quad \sqrt{5+\sqrt{21}},\quad \sqrt{7+2\sqrt{5}},\quad \sqrt{5+2\sqrt{5}}$$

are a few samples. We ask the following question: Is there an expression for α in terms of two square roots which are not nested?

Since $\alpha^2 = r + s\sqrt{t}$, it is easy to write down a quartic polynomial which has α as a root, namely

(6.6) $$f(x) = (x^2 - (r + s\sqrt{t}))(x^2 - (r - s\sqrt{t})) = x^4 + bx^2 + c,$$

where $b = -2r$ and $c = r^2 - s^2t$. If α' denotes one of the two square roots of $r - s\sqrt{t}$, then the roots of this quartic are

(6.7) $$\alpha, \alpha', -\alpha, -\alpha'.$$

The splitting field $K = F(\alpha, \alpha')$ of f can be reached by the sequence $\sqrt{t}, \alpha, \alpha'$ of three square root adjunctions, so the degree $[K : F]$ divides 8. The degree will be less than 8 if one of the square root adjunctions is unnecessary.

We must decide whether or not f is irreducible. To do so, we first check the irreducibility of the quadratic polynomial $q(y) = y^2 + by + c$ whose roots are α^2, α'^2. If q is irreducible, then f doesn't have a root in F. In that case f, if reducible, will be the product of two quadratic polynomials. Computing with undetermined coefficients, we find that the product must have the form

(6.8) $$x^4 + bx^2 + c = (x^2 + ux + v)(x^2 - ux + v).$$

We will be able to determine whether or not such a factorization exists, at least when $F = \mathbb{Q}$.

If $f(x)$ is reducible, then α is a root of a quadratic polynomial, so it can be written using only one square root. This happens with $\sqrt{3+2\sqrt{2}}$ for example, which is equal to $1 + \sqrt{2}$, as you will check by squaring both expressions. The quartics derived from the other examples (6.5) are irreducible over $\mathbb{Q}$.

We now return to our question. Let's suppose that f is irreducible. Notice that to write α in terms of unnested square roots $\sqrt{p}, \sqrt{q}$ amounts to finding a biquadratic extension $K = F(\sqrt{p}, \sqrt{q})$ of F which contains α. Suppose that a biquadratic extension K which contains α can be found. Then K is a Galois extension of F, so $f(x)$ factors into linear factors in K. This means that K contains a splitting field of f. In fact, K will be the splitting field, because f is irreducible and of degree 4. So the Galois group G of f will be the Klein four group. If G is not the Klein four group, then α can not be written in terms of unnested square roots.

Conversely, if K/F is a Galois extension whose Galois group is the Klein four group, then K contains three intermediate fields of degree 2 over F. Any two of these fields taken together generate K. So K is a biquadratic extension of F, and any element of K can be written in terms of two unnested square roots.

We compute the discriminant of $f(x)$, using the list (6.7) of roots.

$$D = \prod_{i<j}(\alpha_i - \alpha_j)^2 = (4\alpha\alpha')^2(\alpha - \alpha')^4(\alpha + \alpha')^4 = 2^4(b^2 - 4c)^2c$$
$$= 2^8 s^4 t^2 (r^2 - s^2 t).$$

If D is a square in F, then G is a transitive subgroup of the alternating group A_4 whose order divides 8. The Klein four group is the only such group. It consists of the even permutations of order 2:

(6.9) $$V = \{(\mathbf{1}), (\mathbf{1\,2})(\mathbf{3\,4}), (\mathbf{1\,3})(\mathbf{2\,4}), (\mathbf{1\,4})(\mathbf{2\,3})\}.$$

There is no other transitive operation of V on $\{\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}\}$. So we find:

(6.10) **Proposition.** Let $\alpha = \sqrt{r+s\sqrt{t}}$, with $r, s, t \in F$, and assume that $f(x) = x^4 - 2rx^2 + (r^2 - s^2t)$ is irreducible over F. Then α can be written in terms of two unnested square roots if and only if $r^2 - s^2t$ is a square in F. □

If $\alpha = \sqrt{5+\sqrt{21}}$, then $r^2 - s^2t = 25 - 21 = 4$, which is a square. In the last two examples (6.5), $r^2 - s^2t$ is not a square in $\mathbb{Q}$.

Let us determine the unnested expression for $\alpha = \sqrt{5+\sqrt{21}}$ explicitly. Galois theory provides the clue; namely it suggests determining the intermediate fields. They are quadratic extensions of $\mathbb{Q}$, so they are generated by square roots. These square roots are the ones we need to express α. One intermediate quadratic extension is obvious, namely $\mathbb{Q}(\sqrt{21})$. But this isn't the one we need. To find another intermediate extension, we determine the fixed field of the subgroup H of order 2 which is generated by $\sigma = (\mathbf{1\,2})(\mathbf{3\,4})$. If the roots of f are listed in the order (6.7), the H-orbit of α is $\{\alpha, \alpha'\}$, (where $\alpha' = \sqrt{5 - \sqrt{21}}$, and the irreducible polynomial for α over K^H is $(x - \alpha)(x - \alpha') = x^2 - (\alpha + \alpha')x + \alpha\alpha'$. So K has degree 2 over the field $L = F(\alpha + \alpha', \alpha\alpha')$, and this field is contained in K^H. A consideration of degrees shows that $L = K^H$. With this clue, we compute, finding $\alpha\alpha' = 2$, $(\alpha + \alpha')^2 = 14$, and $\alpha + \alpha' = \sqrt{14}$. Similarly, $\alpha - \alpha' = \sqrt{6}$. We solve for α, obtaining $\alpha = \frac{1}{2}(\sqrt{6} + \sqrt{14})$. □

It is harder to analyze a general quartic equation, and the roots can usually not be written explicitly in a useful way. However, there is another partially symmetric function which helps to determine the Galois group. Let $f(x)$ be an irreducible quartic polynomial with roots $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ in a splitting field K. Then by Proposition (6.1), its Galois group is a subgroup of S_4, and the roots form one orbit. The transitive subgroups of S_4 are

(6.11) $$S_4,\ A_4,\ D_4,\ C_4,\ V,$$

where V is the group (6.9). Actually, there are three conjugate subgroups isomorphic to D_4 and three conjugate subgroups isomorphic to C_4. The other subgroups are uniquely determined. There are some other subgroups of S_4 which are isomorphic to the Klein four group, but they are not transitive.

Let us ask for partially symmetric functions of the roots to distinguish these groups. As we have seen, the element δ determines whether or not $G \subset A_4$. The subgroups of A_4 in our list are A_4 and V. So $\delta \in F$ if and only if G is one of these two groups.

Next, we consider the partially symmetric polynomial

(6.12) $$\beta_1(u) = u_1u_3 + u_2u_4.$$

A permutation of the indices carries $\beta_1(u)$ to one of the three polynomials $\beta_i(u)$, $i = 1, 2, 3$, where

$$\beta_2(u) = u_1u_2 + u_3u_4 \quad \text{and} \quad \beta_3(u) = u_1u_4 + u_2u_3.$$

Since S_4 has order 24, the stabilizer of $\beta_1(u)$ is of order 8; it is one of the three dihedral groups D_4. The polynomial $(x - \beta_1(u))(x - \beta_2(u))(x - \beta_3(u))$ is left fixed by all permutations of the variables u_i, so its coefficients are symmetric functions. They can be computed explicitly in terms of the elementary symmetric functions.

Going back to our quartic polynomial, we substitute the roots α_i into $\beta_j(u)$, to obtain three elements $\beta_j(\alpha) = \beta_j \in K$. They form one orbit under the action of the symmetric group on the roots. If they are distinct elements of K, then the stabilizer of β_1 in S_4 will have order 8, so it will be the dihedral group D_4. We are lucky: The β_j are distinct. For example,

$$\beta_1 - \beta_2 = \alpha_1\alpha_3 + \alpha_2\alpha_4 - \alpha_1\alpha_2 - \alpha_3\alpha_4 = (\alpha_1 - \alpha_4)(\alpha_3 - \alpha_2).$$

Since we have assumed that f is irreducible, its roots α_i are distinct. The right side of this equation shows that $\beta_1 - \beta_2 \neq 0$.

Since the Galois group G permutes the elements β_i, the polynomial $g(x) = (x-\beta_1)(x-\beta_2)(x-\beta_3)$ has coefficients in F. It is called the *resolvent cubic* of the quartic polynomial $f(x)$.

Though the symmetric group acts transitively on $\{\beta_1, \beta_2, \beta_3\}$, the Galois group G, which is a subgroup of S_4, may not act transitively. Whether or not it does provides information about G. If G fixes β_1 for example, then G is contained in the stabilizer D_4 of β_1. In this case β_1 will be in the field F (1.9), so the resolvent cubic will have a root in F. Proceeding as in the proof of Proposition (6.4), we find the following:

(6.13) **Proposition.** Let $g(x)$ be the resolvent cubic of an irreducible quartic polynomial $f(x)$, and let K be a splitting field of f. Then $g(x)$ has a root in F if and only if the Galois group $G = G(K/F)$ is a subgroup of one of the dihedral groups D_4. In any case, if β is a root of $g(x)$ in K, then the Galois group $G(K/F(\beta))$ is a subgroup of a dihedral group D_4. □

Thus the polynomials $x^2 - D$, where D is the discriminant, and the resolvent cubic $g(x)$ nearly suffice to describe the Galois group. The results are summed up in this table:

(6.14) **Table.**

	D a square in F	D not a square
g reducible	$G = V$	$G = D_4$ or C_4
g irreducible	$G = A_4$	$G = S_4$

Explicit computation for arbitrary quartic equations becomes unpleasant, but we can easily calculate the discriminant of a quartic which has the form

$$x^4 + rx + s. \tag{6.15}$$

The discriminant is a symmetric polynomial of degree 12 and therefore has weighted degree 12 in the elementary symmetric functions $s_1, \ldots, s_4$. Substituting $(0, 0, -r, s)$ for (s_1, s_2, s_3, s_4) into the unknown formula for the discriminant will kill any monomial involving s_1 or s_2. And the only monomials of weighted degree 12 which do not involve s_1 and s_2 are $s_3{}^4$ and $s_4{}^3$. Thus the discriminant of (6.15) has the form

$$D = \Delta(0,0,-r,s) = cr^4 + c's^3.$$

We can determine the coefficients c, c' by computing the discriminant of two particular polynomials. The answer is

$$(6.16) \qquad D = -27r^4 + 256s^3.$$

For example, the discriminant of

$$(6.17) \qquad f(x) = x^4 + 8x + 12$$

is $3^4 \cdot 2^{12}$. This is a square in $\mathbb{Q}$. The Galois group of the splitting field of (6.17) over $\mathbb{Q}$ is therefore a subgroup of A_4.

To calculate the resolvent cubic $g(x)$ of the polynomial (6.15), we write the resolvent cubic for the general polynomial whose roots are $u_1,\dots,u_4$ as

$$g(x) = x^3 - b_1x^2 + b_2x - b_3;$$

then since β_i is a quadratic function in $\{u_j\}$, b_i has degree $2i$ in $\{u_j\}$ and weighted degree $2i$ in the symmetric functions. Proceeding as above, one finds

$$(6.18) \qquad g(x) = x^3 - 4sx - r^2.$$

The resolvent cubic of the particular quartic polynomial (6.17) is $x^3 - 48x - 64$. The quartic (6.17) and its resolvent cubic are both irreducible over $\mathbb{Q}$. It follows that $G = A_4$ for the polynomial (6.17).

7. KUMMER EXTENSIONS

Let us now consider the splitting field over a field F of a polynomial of the form

$$(7.1) \qquad f(x) = x^p - a,$$

where p is a prime. We will assume that the base field F is a subfield of $\mathbb{C}$ which contains the primitive pth root of unity $\zeta_p = e^{2\pi i/p}$. The complex roots of $f(x)$ are the pth roots of a, and if α denotes a particular pth root, then the roots of $f(x)$ are

$$(7.2) \qquad \alpha,\ \zeta\alpha,\ \zeta^2\alpha,\dots,\zeta^{p-1}\alpha,$$

where $\zeta = \zeta_p$. Therefore the splitting field is generated by a single root: $K = F(\alpha)$.

(7.3) **Proposition.** Let F be a subfield of $\mathbb{C}$ which contains the pth root of unity ζ_p, and let a be an element of F which is not a pth power in F. Then the splitting field of $f(x) = x^p - a$ has degree p over F, and its Galois group is a cyclic group of order p.

Proof. Let K be a splitting field of f, and let α be one of its roots in K. Assume that α is not in F. Then there is an automorphism σ of K/F which does not fix

α. Since the roots of f are $\zeta^i\alpha$, $i = 0,\ldots, p-1$, $\sigma(\alpha) = \zeta^\nu\alpha$ for some $\nu \neq 0$. We now compute the powers of σ. Remembering that σ is an automorphism and that $\sigma(\zeta) = \zeta$ because $\zeta \in F$, we find $\sigma^2(\alpha) = \sigma(\zeta^\nu\alpha) = \zeta^\nu\sigma(\alpha) = \zeta^{2\nu}\alpha$. Similarly, $\sigma^i(\alpha) = \zeta^{i\nu}\alpha$ for each i. Since ζ is a pth root of unity, the smallest positive power of σ which fixes α is σ^p. Hence the order of σ in the Galois group is at least p. On the other hand, α generates K over F, and α is a root of the polynomial $x^p - a$ of degree p, so $[K:F] \leq p$. This shows at the same time that $[K:F] = p$, that $x^p - a$ is irreducible over F, and that $G(K/F)$ is cyclic of order p. □

Here is a striking converse to Proposition (7.3):

(7.4) **Theorem.** Let F be a subfield of $\mathbb{C}$ which contains the pth root of unity ζ, and let K/F be a Galois extension of degree p. Then K is obtained by adjoining a pth root to F.

Extensions of this type are often called *Kummer extensions*. For $p = 2$, the theorem reduces to a familiar assertion: Every extension of degree 2 can be obtained by adjoining a square root. But suppose that $p = 3$ and that F contains ζ_3. If the discriminant of the irreducible cubic polynomial (2.3) is a square in F, then the splitting field of f has degree 3 [see (2.16)], so its Galois group is a cyclic group. Therefore the splitting field of such a polynomial has the form $F(\sqrt[3]{a})$, for some $a \in F$. This isn't obvious.

Proof of Theorem (7.4). The Galois group G has prime order $p = [K:F]$, so it is a cyclic group. Any element σ, not the identity, will generate it. Let us view K as an F-vector space. Then σ is a *linear operator* on K. For, since σ is an F-automorphism,

$$\sigma(\alpha+\beta) = \sigma(\alpha) + \sigma(\beta) \quad \text{and} \quad \sigma(c\alpha) = \sigma(c)\sigma(\alpha) = c\sigma(\alpha),$$

for all $c \in F$ and $\alpha, \beta \in K$. Since G is a cyclic group of order p, $\sigma^p = 1$. An eigenvalue λ for this operator must satisfy the relation $\lambda^p = 1$, which means that λ is a power of ζ. By hypothesis, these eigenvalues are in the field F. Moreover, there is at least one eigenvalue different from 1. This is a fact about any linear operator T such that some power of T is the identity, because such a linear operator can be diagonalized [Chapter 9 (2.3)]. Its eigenvalues are the entries of the diagonal matrix A which represents it. If T is not the identity, as is the case here, then $A \neq I$, so some diagonal entry is different from 1.

We choose an eigenvector α with an eigenvalue $\zeta^i \neq 1$. Then $\sigma(\alpha) = \zeta^i\alpha$, and hence $\sigma(\alpha^p) = \sigma(\alpha)^p = (\zeta^i\alpha)^p = \zeta^{ip}\alpha^p = \alpha^p$. So σ fixes α^p. Since σ generates G, the element α^p is in the fixed field K^G, which is F (1.9). We have therefore found an element $\alpha \in K$ whose pth power is in F. Since $\sigma(\alpha) \neq \alpha$, the element α is not in F itself. Since $[K:F]$ is prime, α generates K. □

(7.5) **Example.** Consider the cyclic cubic polynomial (2.12) $x^3 - 3x + 1$. Let $\{\eta_i, \eta_2, \eta_3\}$ denote its roots. There is an element $\sigma \in G(K/F)$ acting as a cyclic

permutation. We choose the basis $(1, \eta_1, \eta_2)$ for K over $F = \mathbb{Q}(\zeta_3)$. (Why is it a basis?) With respect to this basis, the matrix of the linear operator σ is

$$\sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix},$$

because $\sigma(1) = 1$, $\sigma(\eta_1) = \eta_2$, $\sigma(\eta_2) = \eta_3 = -\eta_1 - \eta_2$. The vector $(0, 1, -\zeta_3)^t$ is an eigenvector with eigenvalue ζ_3. Thus if $\alpha = \eta_1 - \zeta_3\eta_2$, then α^3 is an element of F, and α generates the splitting field of $x^3 - 3x + 1$ over F. We can compute α^3 explicitly, using the fact that $\eta_1 = \zeta_9 + \zeta_9{}^8$ and $\eta_2 = \zeta_9{}^2 + \zeta_9{}^7$. Noting that $\zeta_3 = \zeta_9{}^3$, we find $\alpha = \zeta_9{}^8 - \zeta_9{}^5$ and $\alpha^3 = 3(1 - \zeta_3)$. □

(7.6) **Example.** Let $f(x)$ be an arbitrary irreducible cubic polynomial over a field F, and let K be a splitting field of $f(x)(x^3 - 1)$ over F. Let $L \subset K$ be the intermediate field generated by ζ and $\delta = \sqrt{D}$, where D is the discriminant of f. Then $[L : F]$ divides 4, and $[K : L] = 3$, by (2.16). The four elements $\{1, \sqrt{D}, \sqrt{-3}, \sqrt{(-3D)}\}$ span L as F-vector space in any case. By Theorem (7.4), $K = L(\sqrt[3]{b})$, for some $b \in L$. Therefore the roots of $f(x)$ admit some expression in terms of a cube root of the form

$$\sqrt[3]{c_1 + c_2\sqrt{D} + c_3\sqrt{-3} + c_4\sqrt{-3D}}, \quad \text{with } c_i \in F. \;\square$$

8. CYCLOTOMIC EXTENSIONS

The subfield K of the complex numbers which is generated over $\mathbb{Q}$ by $\zeta_n = e^{2\pi i/n}$ is called a *cyclotomic field*. Also, for any subfield F of $\mathbb{C}$, the field $F(\zeta_n)$ is called a *cyclotomic extension* of F. It is the splitting field over F of the polynomial

(8.1) $$x^n - 1.$$

If we denote ζ_n by ζ, the roots of this polynomial are the powers of ζ, the nth roots of unity $1, \zeta, \zeta^2, \ldots, \zeta^{n-1}$. We will concentrate on the case that n is a prime integer p different from 2 in this section.

The polynomial $x^{p-1} + \cdots + x + 1$ is irreducible over $\mathbb{Q}$, and $\zeta = \zeta_p$ is one of its roots [Chapter 11 (4.6)]. So it is the irreducible polynomial for ζ over $\mathbb{Q}$. Its roots are the powers $\zeta, \zeta^2, \ldots, \zeta^{p-1}$. Hence the Galois group of $\mathbb{Q}(\zeta)$ over $\mathbb{Q}$ has order $p - 1$.

(8.2) **Proposition.** Let p be a prime integer, and let $\zeta = \zeta_p$.

(a) The Galois group of $\mathbb{Q}(\zeta)$ over $\mathbb{Q}$ is isomorphic to the multiplicative group $\mathbb{F}_p{}^{\times}$ of nonzero elements of the prime field $\mathbb{F}_p$. It is a cyclic group of order $p - 1$.

(b) For any subfield F of $\mathbb{C}$, the Galois group of $F(\zeta)$ over F is a cyclic group.

Proof. Let G be the Galois group of $F(\zeta)$ over F. We define a map $v: G \longrightarrow \mathbb{F}_p^{\times}$ as follows: Let $\sigma \in G$ be an automorphism. It will carry ζ to another root of the polynomial $x^p + \cdots + x + 1$, say to ζ^i. The exponent i is determined as an integer modulo p, because ζ has multiplicative order p. We set $v(\sigma) = i$. Let us verify that v is multiplicative: If τ is another element of G such that $v(\tau) = j$, that is, $\tau(\zeta) = \zeta^i$, then

$$\sigma\tau(\zeta) = \sigma(\zeta^j) = \sigma(\zeta)^j = \zeta^{ij}. \tag{8.3}$$

Also, the identity automorphism sends ζ to ζ, and hence $v(1) = 1$. Since v is compatible with multiplication and $v(\sigma) \neq 0$, v is a homomorphism to $\mathbb{F}_p^{\times}$. The homomorphism is injective because, since ζ generates K, the action of an automorphism is determined when we know its action on ζ. Thus G is isomorphic to its image in $\mathbb{F}_p^{\times}$. Since $\mathbb{F}_p^{\times}$ is a cyclic group, so is every subgroup. Therefore G is cyclic. If $F = \mathbb{Q}$, then $|G| = |\mathbb{F}_p^{\times}| = p - 1$, so these two groups are isomorphic. □

Suppose that $F = \mathbb{Q}$. Then being cyclic and of order $p - 1$, the Galois group G of $K = \mathbb{Q}(\zeta_p)$ has exactly one subgroup of order k for each integer k which divides $p - 1$. If $(p - 1)/k = r$ and if σ is a generator for G, then the subgroup of order k is generated by σ^r. So by the Main Theorem of Galois theory, there will be exactly one intermediate field L with $[L:\mathbb{Q}] = r$. These fields are generated by certain sums of powers of $\zeta = \zeta_p$. We will illustrate this by some simple examples.

The simplest case is $p = 5$. Then $[K:\mathbb{Q}] = 4$, and there is an intermediate field of degree 2 over $\mathbb{Q}$. It is generated by $\eta = \zeta + \zeta^4 = 2\cos 2\pi/5$. Since $2\cos 2\pi/5 = \frac{1}{2}(-1 + \sqrt{5})$, the intermediate field is the quadratic number field $\mathbb{Q}(\sqrt{5})$.

(8.4) **Proposition.** The subfield L of $K = \mathbb{Q}(\zeta_p)$ whose degree over $\mathbb{Q}$ is $\frac{1}{2}(p - 1)$ is generated over $\mathbb{Q}$ by the element $\eta = \zeta + \zeta^{p-1} = 2\cos 2\pi/p$. Moreover, $L = K \cap \mathbb{R}$.

Since $L = K \cap \mathbb{R}$, L is also called the *real subfield* of K.

Proof. Notice that ζ is a root of the quadratic equation $x^2 - \eta x + 1$, which has coefficients in $\mathbb{Q}(\eta)$. Therefore $[K:\mathbb{Q}(\eta)] \leq 2$. On the other hand, η is a real number, while ζ is not real, so $\mathbb{Q}(\eta) < K$. It follows that $[K : \mathbb{Q}(\eta)] = 2$, that $\mathbb{Q}(\eta) = K \cap \mathbb{R}$, and that $[\mathbb{Q}(\eta) : \mathbb{Q}] = \frac{1}{2}(p - 1)$. □

When $p = 7$, $\eta = \zeta + \zeta^6$ has degree 3 over $\mathbb{Q}$. Its irreducible polynomial over $\mathbb{Q}$ can be computed by a method which we have used before (2.12). We guess that the other roots are $\eta_2 = \zeta^2 + \zeta^5$ and $\eta_3 = \zeta^3 + \zeta^4$. These are the other sums of a pth root and its inverse. It is not hard to show that $\{\eta_1, \eta_2, \eta_3\}$ is the G-orbit of $\eta = \eta_1$, so this guess can be justified formally. We expand $(x - \eta_1)(x - \eta_2)(x - \eta_3)$ and use the relation $\zeta^6 + \cdots + \zeta + 1 = 0$, obtaining the irreducible equation $x^3 + x^2 - 2x - 1$ for η over $\mathbb{Q}$.

The cyclotomic field $\mathbb{Q}(\zeta_7)$ also contains a quadratic extension of $\mathbb{Q}$. It is generated by $\epsilon = \zeta + \zeta^2 + \zeta^4$. If we set $\epsilon' = \zeta^3 + \zeta^5 + \zeta^6$, then $(x - \epsilon)(x - \epsilon') = x^2 + x + 2$ is its irreducible equation. The discriminant of this polynomial is -7, so $\mathbb{Q}(\epsilon) = \mathbb{Q}(\sqrt{-7})$. It follows that $\mathbb{Q}(\zeta_7)$ contains $\sqrt{-7}$.

Suppose that $p = 17$. Then $[\mathbb{Q}(\zeta) : \mathbb{Q}] = 16$. A cyclic group of order 16 contains a chain of subgroups $C_{16} \supset C_8 \supset C_4 \supset C_2 \supset C_1$. By the Main Theorem of Galois theory, there is a corresponding chain of intermediate fields $\mathbb{Q} \subset F_1 \subset F_2 \subset F_3 \subset \mathbb{Q}(\zeta)$, of degrees $1, 2, 4, 8, 16$ over $\mathbb{Q}$. The field F_3 of degree 8 is the real subfield generated by $\eta = 2 \cos 2\pi/17$, as in Proposition (8.4). Since each extension in this chain has degree 2, F_3 can be reached by a succession of three square root adjunctions. This proves that $2 \cos 2\pi/17$, and hence the regular 17-gon, can be constructed by ruler and compass [Chapter 13 (4.9)].

The other field extension which we will describe for all primes is the one of degree 2 over $\mathbb{Q}$. The Main Theorem of Galois theory tells us that there is a unique intermediate field L of $\mathbb{Q}$ of degree 2, corresponding to the subgroup H of G of order $\frac{1}{2}(p - 1)$. If σ generates G, then H is generated by σ^2.

(8.5) **Theorem.** Let p be an odd prime, and let L be the unique quadratic extension of $\mathbb{Q}$ contained in the cyclotomic field $\mathbb{Q}(\zeta_p)$. Then

$$L = \mathbb{Q}(\sqrt{\pm p}),$$

where the sign is $(-1)^{1/2(p-1)}$.

Proof. We need to select a generator of L whose equation is easy to determine. Gauss's method is to take the sum of half of the powers of ζ, suitably chosen.

There is another choice of generator for L which is a little simpler to work with. Let D be the discriminant of the polynomial

$$x^p - 1. \tag{8.6}$$

This discriminant can be computed directly in terms of the roots $\{1, \zeta, \zeta^2, \ldots, \zeta^{p-1}\}$, but it is easier to determine D using the following nice formula:

(8.7) **Lemma.** Let $f(x) = (x - \alpha_1) \cdots (x - \alpha_n)$. The discriminant of f is

$$D = \pm f'(\alpha_1) \cdots f'(\alpha_n) = \pm \prod_i f'(\alpha_i),$$

where f' is the derivative.

Proof. By the product rule for differentiation,

$$f'(x) = \sum_{i=1}^{n} (x - \alpha_1) \cdots (x - \alpha_{i-1})(x - \alpha_{i+1}) \cdots (x - \alpha_n).$$

Therefore

$$f'(\alpha_i) = (\alpha_i - \alpha_1) \cdots (\alpha_i - \alpha_{i-1})(\alpha_i - \alpha_{i+1}) \cdots (\alpha_i - \alpha_n).$$

This is the product of the differences $(\alpha_i - \alpha_j)$, with the given i and with $j \neq i$.

Thus

$$\prod_i f'(\alpha_i) = \prod_{i \neq j} (\alpha_i - \alpha_j) = \pm D. \quad \square$$

We apply this lemma to our polynomial $x^p - 1$. Its derivative is px^{p-1}, so the discriminant is

$$D = \pm \prod_i p\zeta^{i(p-1)} = \pm \zeta^N p^p,$$

where the exponent N is some integer. To determine ζ^N, we note that D is a rational number, because the coefficients of $x^p - 1$ are rational. The only power of ζ which is rational is 1. Therefore $\zeta^N = 1$ and

(8.8)
$$D = \pm p^p.$$

The square root of this discriminant is $\delta = \sqrt{\pm p^p}$. It is in the field $\mathbb{Q}(\zeta)$. Since p is odd and since square factors can be pulled out of a square root,

(8.9)
$$\mathbb{Q}(\delta) = \mathbb{Q}(\sqrt{\pm p}).$$

Therefore this field is a quadratic subfield of $\mathbb{Q}(\zeta)$, and since L is the only quadratic subfield, it is L. We leave the determination of the sign as an exercise. □

The following theorem, first stated by Kronecker, is one of the most beautiful theorems of algebraic number theory. Unfortunately, it would take too long to prove it here.

(8.10) **Theorem.** Every Galois extension K of $\mathbb{Q}$ whose Galois group is abelian is contained in one of the cyclotomic fields $\mathbb{Q}(\zeta_n)$. □

9. QUINTIC EQUATIONS

The main motivation behind Galois' work was the problem of solving fifth-degree equations. We are going to study his solution in this section. A short time earlier, Abel had shown that the quintic

(9.1)
$$x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1 + a_0$$

with variable coefficients a_i could not be solved in terms of radicals, but it remained to find an explicit polynomial with rational coefficients which couldn't be solved. Anyhow, because the problem was over 200 years old, interest in it continued. In the meantime, Galois' ideas have turned out to be much more important than the question which motivated them.

An expression in terms of radicals may become very complicated, and I don't know a good notation for a general one. However, it is easy to give a precise recursive definition. Let F be an arbitrary subfield of the complex numbers. We say that a

complex number α is *expressible by radicals over* F if there is a tower of subfields $F = F_0 \subset F_1 \subset \ldots \subset F_r$ of $\mathbb{C}$ such that

(9.2)

(i) $\alpha \in F_r$, and

(ii) for every $j = 1,\ldots,r$, F_j is generated over F_{j-1} by a radical β_j. In other words, $F_j = F_{j-1}(\beta_j)$, and for some integer n_j, $\beta_j^{n_j} \in F_{j-1}$.

This definition is formally similar to the description [Chapter 13 (4.9)] of the real numbers which can be constructed by ruler and compass. In that description, only square roots of positive real numbers are allowed.

(9.3) **Proposition.** Let α be a root of a polynomial $f(x) \in F[x]$ of degree ≤ 4. Then α is expressible by radicals over F.

Proof. For quadratic polynomials, this is the quadratic formula. For cubics, Cardano's formula gives the solution. Suppose that $f(x)$ is quartic. If f is reducible, then α is a root of a polynomial of lower degree, and the problem is solved. If not, then f has distinct roots in a splitting field K, so its discriminant D is not zero. Let $g(x)$ be the resolvent cubic of f. We proceed by adjoining the square root δ of D, obtaining a field F_1 (possibly equal to F). Next, we use Cardano's formula to solve the resolvent cubic. This will require a square root extension F_2 followed by a cube root extension F_3. At this point, Table (6.14) shows that the Galois group of K/F_3 is a subgroup of the Klein four group. Therefore K can be reached by a sequence of at most two more square root extensions $F_3 \subset F_4 \subset F_5 = K$. □

The nth roots of unity $\zeta_n = e^{2\pi i/n}$ are allowable in an expression by radicals. Also, if $n = rs$, then $\sqrt[n]{b} = \sqrt[r]{\sqrt[s]{b}}$. So at the cost of adding more steps to the chain of fields, we may assume that all the roots are pth roots, for various prime integers p.

Note that there is a great deal of ambiguity in an expression by radicals, because there are n choices for each $\sqrt[n]{b}$. The notation $(-3 + \sqrt[5]{2})^{1/4}$ may stand for any one of 20 complex numbers, so the tower of fields $\mathbb{Q} \subset \mathbb{Q}(\sqrt[5]{2}) \subset \mathbb{Q}((-3 + \sqrt[5]{2})^{1/4})$ is not uniquely defined. This ambiguity is inherent in the notation. Since the notation is cumbersome anyhow, we won't bother trying to make it more precise. We won't use it very much.

(9.4) **Proposition.** Let $f(x)$ be an irreducible polynomial over a field F. If one root of f in K can be expressed by radicals, so can any other root.

Proof. Suppose that one root α can be expressed by radicals, say using the tower $F = F_0 \subset \ldots \subset F_r$. Choose a field L which contains F_r and which is a splitting field of some polynomial of the form $f(x)g(x)$ over F. Then L is also the splitting field of fg over $F(\alpha)$. Let α' be a root of f in another field K', and let L' be a

splitting field of fg over $F(\alpha')$. Then we can extend the isomorphism $F(\alpha) \longrightarrow F(\alpha')$ to an isomorphism $\varphi : L \longrightarrow L'$ (5.2). The tower of fields $F = \varphi(F_0) \subset \ldots \subset \varphi(F_r)$ shows that α' is expressible by radicals. □

(9.5) **Proposition.** Let α be a complex number which can be expressed by radicals over F. Then a tower of fields $F = F_0 \subset \ldots \subset F_r = K$ can be found so that the conditions (i) and (ii) of (9.2) hold and, in addition,

(iii) for each j, F_j is a Galois extension of F_{j-1} and the Galois group $G(F_j/F_{j-1})$ is a cyclic group.

Proof. Consider the tower given in the definition (9.2), in which $F_r = F(\beta_1,\ldots,\beta_r)$. As we have remarked, we may assume that $\beta_j^{p_j} \in F_{j-1}$ for some prime integer p_j. Let $\zeta_{p_j} = e^{2\pi i/p_j}$ be the p_j-th root of 1. We form a new chain of fields by adjoining the elements $(\zeta_{p_1},\ldots,\zeta_{p_r};\beta_1,\ldots,\beta_r)$ in that order. Theorem (7.4) and Proposition (8.2) show that each of these extensions is Galois, with cyclic Galois group. Some of the extensions in this tower may be trivial because of redundancy. If so, we shorten the chain. Since the last field $F(\{\zeta_{p_j}\},\{\beta_j\})$ in this chain contains F_r, it contains α. □

Let us consider the Galois group of a product of polynomials $f(x)g(x)$ over F. Let K' be a splitting field of fg. Then K' contains a splitting field K of f, because f factors into linear factors in K'. Similarly, K' contains a splitting field F' of g. So we have a diagram of fields

(9.6)
$$\begin{array}{ccc} & K' & \\ \cup & & \supset \\ K & & F' \\ \supset & & \cup \\ & F & \end{array}$$

(9.7) **Proposition.** With the above notation, let $G = G(K/F)$, $H = G(F'/F)$, and $\mathcal{G} = G(K'/F)$.

(a) G and H are quotients of $\mathcal{G}$.

(b) $\mathcal{G}$ is isomorphic to a subgroup of the product group $G \times H$.

Proof. The first assertion follows from the fact that K and F' are intermediate fields which are Galois extensions of F (5.7b). Let us denote the canonical homomorphisms $\mathcal{G} \longrightarrow G$, $\mathcal{G} \longrightarrow H$ by subscripts: $\sigma \rightsquigarrow \sigma_f$ and $\sigma \rightsquigarrow \sigma_g$. Then σ_f describes the way that σ operates on the roots of f, and σ_g describes the way it operates on the roots of g. We map $\mathcal{G}$ to $G \times H$ by $\sigma \rightsquigarrow (\sigma_f, \sigma_g)$. If σ_f and σ_g are both the identity, then σ operates trivially on the roots of fg, and hence $\sigma = 1$. This shows that the map $\mathcal{G} \longrightarrow G \times H$ is injective and that $\mathcal{G}$ is isomorphic to a subgroup of $G \times H$. □

(9.8) **Proposition.** Let f be a polynomial over F whose Galois group G is a simple nonabelian group. Let F' be a Galois extension of F, with abelian Galois group. Let K' be a splitting field of f over F'. Then the Galois group $G(K'/F')$ is isomorphic to G.

This proposition is a key point. It tells us that if the Galois group of f is a simple nonabelian group, then we will not make any progress toward solving for its roots if we replace F by an abelian extension F'.

Proof of Proposition (9.8). We first reduce ourselves to the case that $[F' : F]$ is a prime number. To do this, we suppose that the lemma has been proved in that case, and we choose a cyclic quotient group H of $G(F'/F)$ of prime order. Such a quotient exists because $G(F'/F)$ is abelian. This quotient determines an intermediate field $F_1 \subset F'$ which is a Galois extension of F, and such that $G(F_1/F) = H$ (5.7). Let K_1 be the splitting field of f over F_1. Then since $[F_1 : F]$ is a prime, $G(K_1/F_1) = G$. So we may replace F by F_1 and K by K_1. Induction on $[F' : F]$ will complete the proof.

So we may assume that $[F' : F] = p$ and that $H = G(F'/F)$ is a cyclic group of order p. The splitting field K' will contain a splitting field of f over F, call it K. We are then in the situation of Proposition (9.7). So the Galois group $\mathcal{G}$ of K' over F is a subgroup of $G \times H$, and it maps surjectively to G. It follows that $|G|$ divides $|\mathcal{G}|$, and $|\mathcal{G}|$ divides $|G \times H| = p|G|$. If $|G| = |\mathcal{G}|$, then counting degrees shows that $K' = K$. In this case, K contains the Galois extension F', and hence H is a quotient of G (5.7b). Since G is a nonabelian simple group, this is impossible. The only remaining possibility is that $\mathcal{G} = G \times H$. Applying the Main Theorem to the chain of fields $F \subset F' \subset K'$, we conclude that $G(K'/F') = G$, as required. □

(9.9) **Theorem.** The roots of a quintic polynomial $f(x)$ whose Galois group is S_5 or A_5 can not be expressed by radicals over F.

Proof. Let K be a splitting field of f. If $G = S_5$, then the discriminant of f is not a square in F. In that case, we replace F by $F(\delta)$, where δ is a square root of the discriminant in K. The Galois group $G(K/F(\delta))$ is A_5. Obviously, it is enough to show that the roots of f can not be expressed by radicals over the larger field $F(\delta)$. This reduces the case that the group is S_5 to the case that it is A_5.

Suppose that the Galois group of f is A_5 but that some root α of f is expressible by radicals over F. Say that $\alpha \in F_r$, where F_r is the end of a chain of field extensions $F = F_0 \subset \dots \subset F_r$, each extension in the chain being Galois, with a cyclic Galois group. Now since the Galois group of f over F is a simple group, Proposition (9.8) shows inductively that for each i, the Galois group of f over F_i is A_5 too. On the other hand, since it has a root α in F_r, the polynomial f will not remain irreducible over that field. Therefore the Galois group of f over F_r will not operate transitively on the five roots of f in a splitting field. In particular, the Galois group can not be the alternating group. This is a contradiction, which shows that the roots of f are not expressible by radicals. □

We will now exhibit a specific quintic polynomial over $\mathbb{Q}$ whose Galois group is S_5. The facts that 5 is prime and that the Galois group G acts transitively on the roots $\{\alpha_1, \dots \alpha_5\}$ limit the possible Galois groups greatly. For, since the action is transitive, $|G|$ is divisible by 5. Thus G contains an element of order 5. The only elements of order 5 in S_5 are cyclic permutations such as $\sigma = (\mathbf{12345})$.

(9.10) **Lemma.** If G contains a transposition, then $G = S_5$.

Proof. By transposition τ we mean, as always, a permutation which interchanges two indices. We may assume that G contains the cyclic permutation σ above. Renumbering if necessary, we may assume that τ acts as $(\mathbf{1\,i})$. We replace σ by σ^{i-1} and renumber again, to reduce to the case that τ is the transposition $(\mathbf{1\,2})$. It remains only to verify that σ and τ generate S_5, which is left as an exercise. □

(9.11) **Corollary.** Suppose that the irreducible polynomial (9.1) has roots $\{\alpha_1, \dots, \alpha_5\}$, and let K be its splitting field. If $F(\alpha_1, \alpha_2, \alpha_3) < K$, then $G(K/F)$ is the symmetric group S_5.

For let $F' = F(\alpha_1, \alpha_2, \alpha_3)$. The only nontrivial permutation fixing $\alpha_1, \alpha_2, \alpha_3$ is the transposition $(\mathbf{4\,5})$. If $F' \neq K$, this permutation must be in $G(K/F')$. Thus $G(K/F)$ contains a transposition. □

(9.12) **Corollary.** Let $f(x)$ be an irreducible quintic polynomial over $\mathbb{Q}$ with exactly three real roots. Then its Galois group is the symmetric group, and hence its roots can not be expressed by radicals.

For, call the real roots $\alpha_1, \alpha_2, \alpha_3$. Then $\mathbb{Q}(\alpha_1, \alpha_2, \alpha_3) \subset \mathbb{R}$, but since α_4, α_5 are not real, K is not a subfield of $\mathbb{R}$. So we can apply Corollary (9.11) to conclude that the Galois group of f is S_5. By Theorem (9.9), the roots of f can not be expressed by radicals. □

(9.13) **Example.** The polynomial $x^5 - 16x = x(x^2 - 4)(x^2 + 4)$ has three real roots, but of course it is not irreducible. But we can add a small constant without changing the number of real roots. This is seen by looking at the graph of the polynomial. For instance,

$$x^5 - 16x + 2$$

still has three real roots, and it is irreducible by the Eisenstein Criterion [Chapter 10 (4.9)]. So its roots can not be expressed by radicals over $\mathbb{Q}$.

Il parait après cela qu'il n'y a aucun fruit à tirer de la solution que nous proposons.

Evariste Galois

EXERCISES

1. The Main Theorem of Galois Theory

1. Determine the irreducible polynomial for $i + \sqrt{2}$ over $\mathbb{Q}$.
2. Prove that the set $(1, i, \sqrt{2}, i\sqrt{2})$ is a basis for $\mathbb{Q}(i, \sqrt{2})$ over $\mathbb{Q}$.
3. Determine the intermediate fields between $\mathbb{Q}$ and $\mathbb{Q}(\sqrt{2}, \sqrt{3})$.
4. Determine the intermediate fields of an arbitrary biquadratic extension without appealing to the Main Theorem.
5. Prove that the automorphism $\mathbb{Q}(\sqrt{2})$ sending $\sqrt{2}$ to $-\sqrt{2}$ is discontinuous.
6. Determine the degree of the splitting field of the following polynomials over $\mathbb{Q}$.
   (a) $x^4 - 1$ **(b)** $x^3 - 2$ **(c)** $x^4 + 1$
7. Let α denote the positive real fourth root of 2. Factor the polynomial $x^4 - 2$ into irreducible factors over each of the fields $\mathbb{Q}$, $\mathbb{Q}(\sqrt{2})$, $\mathbb{Q}(\sqrt{2}, i)$, $\mathbb{Q}(\alpha)$, $\mathbb{Q}(\alpha, i)$.
8. Let $\zeta = e^{2\pi i/5}$.
   (a) Prove that $K = \mathbb{Q}(\zeta)$ is a splitting field for the polynomial $x^5 - 1$ over $\mathbb{Q}$, and determine the degree $[K : \mathbb{Q}]$.
   (b) Without using Theorem (1.11), prove that K is a Galois extension of $\mathbb{Q}$, and determine its Galois group.
9. Let K be a quadratic extension of the form $F(\alpha)$, where $\alpha^2 = a \in F$. Determine all elements of K whose squares are in F.
10. Let $K = \mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{5})$. Determine $[K : \mathbb{Q}]$, prove that K is a Galois extension of $\mathbb{Q}$, and determine its Galois group.
11. Let K be the splitting field over $\mathbb{Q}$ of the polynomial $f(x) = (x^2 - 2x - 1)(x^2 - 2x - 7)$. Determine $G(K/\mathbb{Q})$, and determine all intermediate fields explicitly.
12. Determine all automorphisms of the field $\mathbb{Q}(\sqrt[3]{2})$.
13. Let K/F be a finite extension. Prove that the Galois group $G(K/F)$ is a finite group.
14. Determine all the quadratic number fields $\mathbb{Q}[\sqrt{d}]$ which contain a primitive pth root of unity, for some prime $p \neq 2$.
15. Prove that every Galois extension K/F whose Galois group is the Klein four group is biquadratic.
16. Prove or disprove: Let $f(x)$ be an irreducible cubic polynomial in $\mathbb{Q}[x]$ with one real root α. The other roots form a complex conjugate pair $\beta, \overline{\beta}$, so the field $L = \mathbb{Q}(\beta)$ has an automorphism σ which interchanges $\beta, \overline{\beta}$.
17. Let K be a Galois extension of a field F such that $G(K/F) \approx C_2 \times C_{12}$. How many intermediate fields L are there such that **(a)** $[L : F] = 4$, **(b)** $[L : F] = 9$, **(c)** $G(K/L) \approx C_4$?
18. Let $f(x) = x^4 + bx^2 + c \in F[x]$, and let K be the splitting field of f. Prove that $G(K/F)$ is contained in a dihedral group D_4.
19. Let $F = \mathbb{F}_2(u)$ be the rational function field over the field of two elements. Prove that the polynomial $x^2 - u$ is irreducible in $F[x]$ and that it has two equal roots in a splitting field.

20. Let F be a field of characteristic 2, and let K be an extension of F of degree 2.

(a) Prove that K has the form $F(\alpha)$, where α is the root of an irreducible polynomial over F of the form $x^2 + x + a$, and that the other root of this equation is $\alpha + 1$.

(b) Is it true that there is an automorphism of K sending $\alpha \rightsquigarrow \alpha + 1$?

2. Cubic Equations

1. Prove that the discriminant of a real cubic is positive if all the roots are real, and negative if not.

2. Determine the Galois groups of the following polynomials.

(a) $x^3 - 2$ (b) $x^3 + 27x - 4$ (c) $x^3 + x + 1$ (d) $x^3 + 3x + 14$
(e) $x^3 - 3x^2 + 1$ (f) $x^3 - 21x + 7$ (g) $x^3 + x^2 - 2x - 1$
(h) $x^3 + x^2 - 2x + 1$

3. Let f be an irreducible cubic polynomial over F, and let δ be the square root of the discriminant of f. Prove that f remains irreducible over the field $F(\delta)$.

4. Let α be a complex root of the polynomial $x^3 + x + 1$ over $\mathbb{Q}$, and let K be a splitting field of this polynomial over $\mathbb{Q}$.

(a) Is $\sqrt{-3}$ in the field $\mathbb{Q}(\alpha)$? Is it in K?

(b) Prove that the field $\mathbb{Q}(\alpha)$ has no automorphism except the identity.

***5.** Prove Proposition (2.16) directly for a cubic of the form (2.3), by determining the formula which expresses α_2 in terms of α_1, δ, p, q explicitly.

6. Let $f \in \mathbb{Q}[x]$ be an irreducible cubic polynomial which has exactly one real root, and let K be its splitting field over $\mathbb{Q}$. Prove that $[K : \mathbb{Q}] = 6$.

7. When does the polynomial $x^3 + px + q$ have a multiple root?

8. Determine the coefficients p, q which are obtained from the general cubic (2.1) by the substitution (2.2).

9. Prove that the discriminant of the cubic $x^3 + px + q$ is $-4p^3 - 27q^2$.

3. Symmetric Functions

1. Derive the expression (3.10) for the discriminant of a cubic by the method of undetermined coefficients.

2. Let $f(u)$ be a symmetric polynomial of degree d in $u_1, \ldots, u_n$, and let $f^0(u_1, \ldots, u_{n-1}) = f(u_1, \ldots, u_{n-1}, 0)$. Say that $f^0(u) = g(s^0)$, where s_i^0 are the elementary symmetric functions in $u_1, \ldots, u_{n-1}$. Prove that if $n > d$, then $f(u) = g(s)$.

3. Compute the discriminant of a quintic polynomial of the form $x^5 + ax + b$.

4. With each of the following polynomials, determine whether or not it is a symmetric function, and if so, write it in terms of the elementary symmetric functions.

(a) $u_1^2 u_2 + u_2^2 u_1$ $(n = 2)$
(b) $u_1^2 u_2 + u_2^2 u_3 + u_3^2 u_1$ $(n = 3)$
(c) $(u_1 + u_2)(u_2 + u_3)(u_1 + u_3)$ $(n = 3)$
(d) $u_1^3 u_2 + u_2^3 u_3 + u_3^3 u_1 - u_1 u_2^3 - u_2 u_3^3 - u_3 u_1^3$ $(n = 3)$
(e) $u_1^3 + u_2^3 + \cdots + u_n^3$

5. Find two natural bases for the ring of symmetric functions, as free module over the ring R.

*6. Define the polynomials $w_1, \ldots, w_n$ in variables $u_1, \ldots, u_n$ by $w_k = u_1{}^k + \cdots + u_n{}^k$.
 (a) Prove *Newton's identities:* $w_k - s_1 w_{k-1} + s_2 w_{k-2} - \cdots \pm s_{k-1} w_1 \mp k s_k = 0$.
 (b) Do $w_1, \ldots, w_n$ generate the ring of symmetric functions?

7. Let $f(x) = x^3 + a_2x^2 + a_1x + a_0$. Prove that the substitution $x = x_1 - (a_2/3)$ does not change the discriminant of a cubic polynomial.

8. Prove that $[F(u) : F(s)] = n!$ by induction, directly from the definitions.

9. Let $u_1, \ldots, u_n$ be variables and let D_1 denote the discriminant. Define

$$D_2 = \sum_k \prod_{\substack{i<j \\ i,j \neq k}} (u_i - u_j)^2.$$

 (a) Prove that D_2 is a symmetric polynomial, and compute its expression in terms of the elementary symmetric polynomials for the cases $n = 2, 3$.
 (b) Let $a_1, \ldots, a_n$ be elements of a field of characteristic zero. Prove that $D_1(a_1, \ldots, a_n) = D_2(a_1, \ldots, a_n) = 0$ if and only if the number of distinct elements in the set $\{a_1, \ldots, a_n\}$ is $\leq n - 2$.

10. Compute the discriminants of the polynomials given in Section 2, exercise 2.

*11. (*Vandermonde* determinant) (a) Prove that the determinant of the matrix

$$\begin{bmatrix} 1 & u_1 & u_1{}^2 & \cdots & u_1{}^{n-1} \\ 1 & u_2 & & & u_2{}^{n-1} \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ 1 & u_n & u_n{}^2 & \cdots & u_n{}^{n-1} \end{bmatrix}$$

 is a constant multiple of $\delta(u)$.
 (b) Determine the constant.

4. Primitive Elements

1. Let G be a group of automorphisms of a field K. Prove that the fixed elements K^G form a subfield of K.

2. Let $\alpha = \sqrt[3]{2}$, $\zeta = \frac{1}{2}(-1 + \sqrt{-3})$, $\beta = \alpha\zeta$.
 (a) Prove that for all $c \in \mathbb{Q}$, $\gamma = \alpha + c\beta$ is the root of a sixth-degree polynomial of the form $x^6 + ax^3 + b$.
 (b) Prove that the irreducible polynomial for $\alpha + \beta$ is cubic.
 (c) Prove that $\alpha - \beta$ has degree 6 over $\mathbb{Q}$.

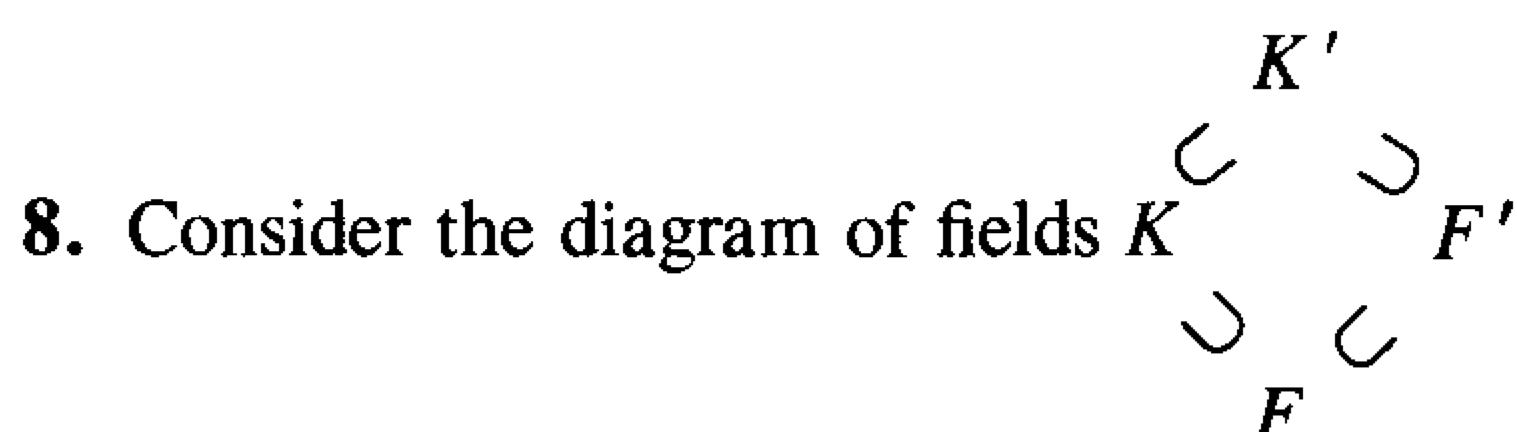
3. For each of the following sets of automorphisms of the field of rational functions $\mathbb{C}(y)$, determine the group of automorphisms which they generate, and determine the fixed field explicitly.
 (a) $\sigma(y) = y^{-1}$ (b) $\sigma(y) = iy$ (c) $\sigma(y) = -y$, $\tau(y) = y^{-1}$ (d) $\sigma(y) = \zeta y$, $\tau(y) = y^{-1}$, where $\zeta = e^{2\pi i/3}$ (e) $\sigma(y) = iy$, $\tau(y) = y^{-1}$

4. (a) Show that the automorphisms $\sigma(y) = (y + i)/(y - i)$, $\tau(y) = i(y - 1)/(y + 1)$ of $\mathbb{C}(y)$ generate a group isomorphic to the alternating group A_4.
 *(b) Determine the fixed field of this group.

*5. Let F be a finite field, and let $f(x)$ be a nonconstant polynomial whose derivative is the zero polynomial. Prove that f is not irreducible over F.

5. *Proof of the Main Theorem*

1. Let $K = \mathbb{Q}(\alpha)$, where α is a root of the polynomial $x^3 + 2x + 1$, and let $g(x) = x^3 + x + 1$. Does $g(x)$ have a root in K?
2. Let $f \in F[x]$ be a polynomial of degree n, and let K be a splitting field for f. Prove that $[K : F]$ divides $n!$.
3. Let G be a finite group. Prove that there exists a field F and a Galois extension K of F whose Galois group is G.
4. Assume it known that π and e are transcendental numbers. Let K be the splitting field of the polynomial $x^3 + \pi x + 6$ over the field $F = \mathbb{Q}(\pi)$.
   (a) Prove that $[K : F] = 6$.
   (b) Prove that K is isomorphic to the splitting field of $x^3 + ex + 6$ over $\mathbb{Q}(e)$.
5. Prove the isomorphism $F[x]/(f(x)) \approx \tilde{F}[x]/(\tilde{f}(x))$ used in the proof of Lemma (5.1) formally, using the universal property of the quotient construction.
6. Prove Corollary (5.5).
7. Let $f(x)$ be an irreducible cubic polynomial over $\mathbb{Q}$ whose Galois group is S_3. Determine the possible Galois groups of the polynomial $(x^3 - 1) \cdot f(x)$.
8. Consider the diagram of fields



   in which K is a Galois extension of F, and K' is generated over F by K and F'. Prove that K' is a Galois extension of F' and that its Galois group is isomorphic to a subgroup of $G(K/F)$.
9. Let $K \supset L \supset F$ be fields. Prove or disprove:
   (a) If K/F is Galois, then K/L is Galois.
   (b) If K/F is Galois, then L/F is Galois.
   (c) If L/F and K/L are Galois, then K/F is Galois.
10. Let K be a splitting field of an irreducible cubic polynomial $f(x)$ over a field F whose Galois group is S_3. Determine the group $G(F(\alpha)/F)$ of automorphisms of the extension $F(\alpha)$.
11. Let K/F be a Galois extension whose Galois group is the symmetric group S_3. Is it true that K is the splitting field of an irreducible cubic polynomial over F?
12. Let K/F be a field extension of characteristic $p \neq 0$, and let α be a root in K of an irreducible polynomial $f(x) = x^p - x - a$ over F.
    (a) Prove that $\alpha + 1$ is also a root of $f(x)$.
    (b) Prove that the Galois group of f over F is cyclic of order p.

6. *Quartic Equations*

1. Compute the discriminant of the quartic polynomial $x^4 + 1$, and determine its Galois group over $\mathbb{Q}$.
2. Let K be the splitting field of an irreducible quartic polynomial $f(x)$ over F, and let the roots of $f(x)$ in K be $\alpha_1, \alpha_2, \alpha_3, \alpha_4$. Also assume that the resolvent cubic $g(x)$ has a root,

say $\beta_1 = \alpha_1\alpha_2 + \alpha_3\alpha_4$. Express the root α_1 explicitly in terms of a succession of square roots.

3. What can you say about the Galois group of an irreducible quartic polynomial over $\mathbb{Q}$ which has exactly two real roots?

4. Suppose that a real quartic polynomial has a positive discriminant. What can you say about the number of real roots?

5. Let K be the splitting field of a reducible quartic polynomial with distinct roots over a field F. What are the possible Galois groups of K/F?

6. What are the possible Galois groups over $\mathbb{Q}$ of an irreducible quartic polynomial $f(x)$ whose discriminant is negative?

7. Let g be the resolvent cubic of an irreducible quartic polynomial $f \in F[x]$. Determine the possible Galois groups of g over F, and in each case, say what you can about the Galois group of f.

8. Let K be the splitting field of a polynomial $f \in F[x]$ with distinct roots $\alpha_1,\dots,\alpha_n$, and let $G = G(K/F)$. Then G may be regarded as a subgroup of the symmetric group S_n. Prove that a change of numbering of the roots changes G to a conjugate subgroup.

9. Let $\alpha_1,\dots,\alpha_4$ be the roots of a quartic polynomial. Discuss the symmetry of the elements $\alpha_1\alpha_2$ and $\alpha_1 + \alpha_2$ along the lines of the discussion in the text.

10. Find a quartic polynomial over $\mathbb{Q}$ whose Galois group is **(a)** S_4, **(b)** D_4, **(c)** C_4.

11. Let α be the real root of a quartic polynomial f over $\mathbb{Q}$. Assume that the resolvent cubic is irreducible. Prove that α can't be constructed by ruler and compass.

12. Determine the Galois groups of the following polynomials over $\mathbb{Q}$.
(a) $x^4 + 4x^2 + 2$ **(b)** $x^4 + 2x^2 + 4$ **(c)** $x^4 + 4x^2 - 5$ **(d)** $x^4 - 2$ **(e)** $x^4 + 2$
(f) $x^4 + 1$ **(g)** $x^4 + x + 1$ **(h)** $x^4 + x^3 + x^2 + x + 1$ **(i)** $x^4 + x^2 + 4$

13. Compute the discriminant of the quartic polynomial $x^4 + ax + b$, using the formula in Lemma (8.7).

***14.** Let f be an irreducible quartic polynomial over F of the form $x^4 + rx + s$, and let $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ be the roots of f in a splitting field K. Let $\eta = \alpha_1\alpha_2$.
(a) Prove that η is the root of a sextic polynomial $h(x)$ with coefficients in F.
(b) Assume that the six products $\alpha_i\alpha_j$ are distinct. Prove that $h(x)$ is irreducible, or else it has an irreducible quadratic factor.
(c) Describe the possibilities for the Galois group $G = G(K/F)$ in the following three cases: h is irreducible, h is a product of an irreducible quadratic and an irreducible quartic, and h is the product of three irreducible quadratics.
(d) Describe the situation when some of the products are equal.

15. Let K be the splitting field of the polynomial $x^4 - 3$ over $\mathbb{Q}$.
(a) Prove that $[K : \mathbb{Q}] = 8$ and that K is generated by i and a single root α of the polynomial.
(b) Prove that the Galois group of $K/\mathbb{Q}$ is dihedral, and describe the operation of the elements of G on the generators of K explicitly.

16. Let K be the splitting field over $\mathbb{Q}$ of the polynomial $x^4 - 2x^2 - 1$. Determine the Galois group G of $K/\mathbb{Q}$, find all intermediate fields, and match them up with the subgroups of G.

17. Let $f(x)$ be a quartic polynomial. Prove that the discriminants of f and of its resolvent cubic are equal.

18. Prove the irreducibility of the polynomial (6.17) and of its resolvent cubic.

19. Let K be the splitting field of the reducible polynomial $(x - 1)^2(x^2 + 1)$ over $\mathbb{Q}$. Prove that $\delta \in \mathbb{Q}$, but that $G(K/\mathbb{Q})$ is not contained in the alternating group.

20. Let $f(x)$ be a quartic polynomial with distinct roots, whose resolvent cubic $g(x)$ splits completely in the field F. What are the possible Galois groups of $f(x)$?

21. Let $\zeta = e^{2\pi i/3}$ be the cube root of 1, let $\alpha = \sqrt[3]{a+b\sqrt{2}}$, and let K be the splitting field of the irreducible polynomial for α over $\mathbb{Q}(\zeta)$. Determine the possible Galois groups of K over $\mathbb{Q}(\zeta)$.

22. Let $\mathcal{H}$ be a subgroup of the symmetric group S_n. Given any monomial m, we can form the polynomial $p(u) = \Sigma_{\sigma \in \mathcal{H}} \sigma m$. Show that if $m = u_1 u_2^2 u_3^3 \cdots u_{n-1}^{n-1}$, then $p(u)$ is partially symmetric for $\mathcal{H}$; that is, it is fixed by the permutations in $\mathcal{H}$ but not by any other permutations.

23. Let $p(u)$ be the polynomial formed as in the last problem, with $\mathcal{H} = A_n$. Then the orbit of $p(u)$ contains two elements, say $p(u), q(u)$. Prove that $p(u) - q(u) = \pm\delta(u)$.

24. Determine the possible Galois groups of a reducible quartic equation of the form $x^4 + bx^2 + c$, assuming that the quadratic $y^2 + by + c$ is irreducible.

25. Compute the discriminant of the polynomial $x^4 + rx + s$ by evaluating the discriminants of $x^4 - x$ and $x^4 - 1$.

26. Use the substitution $x \rightsquigarrow y^{-1}$ to determine the discriminant of the polynomial $x^4 + ax^3 + b$.

27. Determine the resolvent cubic of the polynomials **(a)** $x^4 + rx + s$ and **(b)** $x^4 + a_1x^3 + a_2x^2 + a_3x + a_4$.

28. Let $f(x) = x^4 - 2rx^2 + (r^2 - s^2v)$, with $r, s, v \in F$. Assume that f is irreducible, and let G denote its Galois group. Let $L = F(\sqrt{v}, \delta)$, where $\delta^2 = D$. Prove each statement.
(a) $L(\alpha) = K$
(b) If $[L : F] = 4$, then $G = D_4$.
(c) If $[L : F] = 2$ and $\delta \notin F$, then $G = C_4$.

29. Determine the Galois groups of the last two examples of (6.5).

30. Determine the action of the Galois group G on the roots $\{\alpha, \alpha', -\alpha, -\alpha'\}$ (6.7) explicitly, assuming that **(a)** $G = C_4$, **(b)** $G = D_4$.

31. Determine whether or not the following nested radicals can be written in terms of unnested ones, and if so, find an expression.
(a) $\sqrt{2+\sqrt{11}}$ **(b)** $\sqrt{6+\sqrt{11}}$ **(c)** $\sqrt{11+6\sqrt{2}}$ **(d)** $\sqrt{11+\sqrt{6}}$

***32.** Let K be the splitting field of a quartic polynomial $f(x)$ over $\mathbb{Q}$, whose Galois group is D_4, and let α be a real root of $f(x)$ in K. Decide whether or not α can be constructed by ruler and compass if (a) all four roots of f are real, (b) f has two real roots.

33. Can the roots of the polynomial $x^4 + x - 5$ be constructed by ruler and compass?

7. *Kummer Extensions*

1. Suppose that a Galois extension K/F has the form $K = F(\alpha)$ and that for some integer n, $\alpha^n \in F$. What can you say about the Galois group of K/F?

***2.** Let a be an element of a field F, and let p be a prime. Suppose that the polynomial $x^p - a$ is reducible in $F[x]$. Prove that it has a root in f.

3. Let F be a subfield of $\mathbb{C}$ which contains i, and let K be a Galois extension of F whose group is C_4. Is it true that K has the form $F(\alpha)$, where $\alpha^4 \in F$?

4. Let $f(x) = x^3 + px + q$ be an irreducible polynomial over a field F, with roots $\alpha_1, \alpha_2, \alpha_3$. Let $\beta = \alpha_1 + \zeta\alpha_2 + \zeta^2\alpha_3$, where $\zeta = e^{2\pi i/3}$. Show that β is an eigenvector of σ for the cyclic permutation of the roots unless $\beta = 0$, and compute β^3 explicitly in terms of p, q, δ, ζ.

5. Let K be a splitting field of an irreducible polynomial $f(x) \in F[x]$ of degree p whose Galois group is a cyclic group of order p generated by σ, and suppose that F contains the pth root of unity $\zeta = \zeta_p$. Let $\alpha_1, \alpha_2, \ldots, \alpha_p$ be the roots of f in K. Show that $\beta = \alpha_1 + \zeta^\nu\alpha_2 + \zeta^{2\nu}\alpha_3 + \cdots + \zeta^{(p-1)\nu}\alpha_p$ is an eigenvector of σ, with eigenvalue $\zeta^{-\nu}$, unless it is zero.

6. Let $f(x) = x^3 + px + q$ be an irreducible polynomial over a subfield F of the complex numbers, with complex roots $\alpha = \alpha_1, \alpha_2, \alpha_3$. Let $K = F(\alpha)$.
(a) Express $(6\alpha^2 + 2p)^{-1}$ explicitly, as a polynomial of degree 2 in α.
(b) Assume that $\delta = \sqrt{D}$ is in F, so that K contains the other roots of f. Express α_2 as a polynomial in $\alpha = \alpha_1$ and δ.
(c) Prove that $(1, \alpha_1, \alpha_2)$ is a basis of K, as F-vector space.
(d) Let σ be the automorphism of K which permutes the three roots cyclically. Write the matrix of φ with respect to the above basis, and find its eigenvalues and eigenvectors.
(e) Let v be an eigenvector with eigenvalue $\zeta = e^{2\pi i/3}$. Prove that if $\sqrt{-3} \in F$ then $v^3 \in F$. Compute v^3 explicitly, in terms of $p, q, \delta, \sqrt{-3}$.
(f) Dropping the assumptions that δ and $\sqrt{-3}$ are in F, express v in terms of radicals.
(g) Without calculation, determine the element v' which is obtained from v by interchanging the roles of α_1, α_2.
(h) Express the root α_1 in terms of radicals.

8. Cyclotomic Extensions

1. Determine the degree of ζ_7 over the field $\mathbb{Q}(\zeta_3)$.

2. Let $\zeta = \zeta_{13}$, and let $K = \mathbb{Q}(\zeta)$. Determine the intermediate field of degree 3 over $\mathbb{Q}$ explicitly.

3. Let $\zeta = \zeta_{17}$. Determine the succession of square roots which generate the field $\mathbb{Q}(\zeta + \zeta^{16})$ explicitly.

4. Let $\zeta = \zeta_7$. Determine the degree of the following elements over $\mathbb{Q}$.
(a) $\zeta + \zeta^5$ **(b)** $\zeta^3 + \zeta^4$ **(c)** $\zeta^3 + \zeta^5 + \zeta^6$

5. Let $\zeta = \zeta_{13}$. Determine the degree of the following elements over $\mathbb{Q}$.
(a) $\zeta + \zeta^{12}$ **(b)** $\zeta + \zeta^2$ **(c)** $\zeta + \zeta^5 + \zeta^8$ **(d)** $\zeta^2 + \zeta^5 + \zeta^6$
(e) $\zeta + \zeta^5 + \zeta^8 + \zeta^{12}$ **(f)** $\zeta + \zeta^2 + \zeta^5 + \zeta^{12}$ **(g)** $\zeta + \zeta^3 + \zeta^4 + \zeta^9 + \zeta^{10} + \zeta^{12}$

6. Let $\zeta = \zeta_{11}$.
(a) Prove that $\alpha = \zeta + \zeta^3 + \zeta^4 + \zeta^5 + \zeta^9$ generates a field of degree 2 over $\mathbb{Q}$, and find its equation.
(b) Find an element which generates a subfield of degree 5 over $\mathbb{Q}$, and find its equation.

7. Prove that every quadratic extension of $\mathbb{Q}$ is contained in a cyclotomic extension.

8. Let $K = \mathbb{Q}(\zeta_n)$.
(a) Prove that K is a Galois extension of $\mathbb{Q}$.

(b) Define an injective homomorphism $v: G(K/\mathbb{Q}) \longrightarrow U$ to the group U of units in the ring $\mathbb{Z}/(n)$.

(c) Prove that this homormophism is bijective when $n = 6, 8, 12$. (Actually, this map is always bijective.)

***9.** Let p be a prime, and let a be a rational number which is not a pth power. Let K be a splitting field of the polynomial $x^p - a$ over $\mathbb{Q}$.

(a) Prove that K is generated over $\mathbb{Q}$ by a pth root α of a and a primitive pth root ζ of unity.

(b) Prove that $[K : \mathbb{Q}] = p(p - 1)$.

(c) Prove that the Galois groups of $K/\mathbb{Q}$ is isomorphic to the group of invertible 2×2 matrices with entries in $\mathbb{F}_p$ of the form $\begin{bmatrix} a & b \\ & 1 \end{bmatrix}$, and describe the actions of the elements $\begin{bmatrix} a & \\ & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & b \\ & 1 \end{bmatrix}$ on the generators explicitly.

10. Determine the Galois group of the polynomials $x^8 - 1$, $x^{12} - 1$, $x^9 - 1$.

11. (a) Characterize the primes p such that the regular p-gon can be constructed by ruler and compass.

(b) Extend the characterization to the case of an n-gon, where n is not necessarily prime.

***12.** Let ν be a primitive element modulo a prime p, and let d be a divisor of $p - 1$. Show how to determine a sum of powers of $\zeta = \zeta_p$ which generates the subfield L of $\mathbb{Q}(\zeta)$ of degree d over $\mathbb{Q}$, using the list of roots of unity $\{\zeta, \zeta^\nu, \zeta^{\nu^2}, \ldots, \zeta^{\nu^{p-2}}\}$.

9. Quintic Equations

1. Determine the transitive subgroups of S_5.

2. Let G be the Galois group of an irreducible quintic polynomial. Show that if G contains an element of order 3, then $G = S_5$ or A_5.

***3.** Let p be a prime integer, and let G be a p-group. Let H be a proper normal subgroup of G.

(a) Prove that the normalizer $N(H)$ of H is strictly larger than H.

(b) Prove that H is contained in a subgroup of index p and that that subgroup is normal in G.

(c) Let K be a Galois extension of $\mathbb{Q}$ whose degree is a power of 2, and such that $K \subset \mathbb{R}$. Prove that the elements of K can be constructed by ruler and compass.

4. Let $K \supset L \supset F$ be a tower of field extensions of degree 2. Show that K can be generated over F by the root of an irreducible quartic polynomial of the form $x^4 + bx^2 + c$.

***5.** Cardano's Formula has a peculiar feature: Suppose that the coefficients p, q of the cubic are real numbers. A real cubic always has at least one real root. However, the square root appearing in the formula (2.6) will be imaginary if $(q/2)^2 + (p/3)^3 < 0$. In that case, the real root is displayed in terms of an auxiliary complex number u. This was considered to be an improper solution in Cardano's time. Let $f(x)$ be an irreducible cubic over a subfield F of $\mathbb{R}$, which has three real roots. Prove that no root of f is expressible by real radicals, that is, that there is no tower $F = F_0 \subset \ldots \subset F_r$ as in (9.2), in which all the fields are subfields of $\mathbb{R}$.

6. Let $f(x) \in F[x]$ be an irreducible quintic polynomial, and let K be a splitting field for $f(x)$ over F.

(a) What are the possible Galois groups $G(K/F)$, assuming that the discriminant D is a square in F?

*(b) What are the possible Galois groups if D is not a square in F?

7. Determine which real numbers α of degree 4 over $\mathbb{Q}$ can be constructed with ruler and compass in terms of the Galois group of the corresponding polynomial.

8. Is every Galois extension of degree 10 solvable by radicals?

***9.** Find a polynomial of degree 7 over $\mathbb{Q}$ whose Galois group is S_7.

Miscellaneous Problems

1. Let K be a Galois extension of F whose Galois group is the symmetric group S_4. What numbers occur as degrees of elements of K over F?

2. Show without computation that the side length of a regular pentagon inscribed in the unit circle has degree 2 over $\mathbb{Q}$.

3. (a) The nonnegative real numbers are those having a real square root. Use this fact to prove that the field $\mathbb{R}$ has no automorphism except the identity.

(b) Prove that $\mathbb{C}$ has no *continuous* automorphisms except for complex conjugation and the identity.

4. Let K/F be a Galois extension with Galois group G, and let H be a subgroup of G. Prove that there exists an element $\beta \in K$ whose stabilizer is H.

***5.** (a) Let K be a field of characteristic p. Prove that the *Frobenius* map φ defined by $\varphi(x) = x^p$ is a homomorphism from K to itself.

(b) Prove that φ is an isomorphism if K is a finite field.

(c) Give an example of an infinite field of characteristic p such that φ is not an isomorphism.

(d) Let $K = \mathbb{F}_q$, where $q = p^r$, and let $F = \mathbb{F}_p$. Prove that $G(K/F)$ is a cyclic group of order r, generated by the Frobenius map φ.

(e) Prove that the Main Theorem of Galois theory holds for the field extension K/F.

6. Let K be a subfield of $\mathbb{C}$, and let G be its group of automorphisms. We can view G as acting on the point set K in the complex plane. The action will probably be discontinuous, but nevertheless, we can define an action on line segments $[\alpha, \beta]$ whose endpoints are in K, by defining $g[\alpha, \beta] = [g\alpha, g\beta]$. Then G also acts on polygons whose vertices are in K.

(a) Let $K = \mathbb{Q}(\zeta)$, where ζ is a primitive fifth root of 1. Find the G-orbit of the regular pentagon whose vertices are $1, \zeta, \zeta^2, \zeta^3, \zeta^4$.

(b) Let α be the side length of the pentagon of (a). Show that $a = \alpha^2 \in K$, and find the irreducible equation for α over $\mathbb{Q}$. Is $\alpha \in K$?

7. A polynomial $f \in F[x_1, \ldots, x_n]$ is called $\frac{1}{2}$-symmetric if $f(u_{\sigma 1}, \ldots, u_{\sigma n}) = f(u_1, \ldots, u_n)$ for every even permutation σ of the indices, and skew-symmetric if $f(u_{\sigma 1}, \ldots, u_{\sigma n}) =$ (sign $\sigma) f(u_1, \ldots, u_n)$ for every permutation σ.

(a) Prove that the square root of the discriminant $\delta = \Pi_{i<j} (u_i - u_j)$ is skew-symmetric.

(b) Prove that every $\frac{1}{2}$-symmetric polynomial has the form $f + g\delta$, where f, g are symmetric polynomials.

***8.** Let $f(x, y) \in \mathbb{C}[x, y]$ be an irreducible polynomial, which we regard as a polynomial $f(y)$ in y. Assume that f is cubic as a polynomial in y. Its discriminant D, computed

with regard to the variable y, will be a polynomial in x. Assume that there is a root x_0 of $D(x)$ which is not a multiple root.

(a) Prove that the polynomial $f(x_0, y)$ in y has one simple root and one double root.

(b) Prove that the splitting field K of $f(y)$ over $\mathbb{C}(x)$ has degree 6.

9. Let K be a subfield of $\mathbb{C}$ which is a Galois extension of $\mathbb{Q}$. Prove or disprove: Complex conjugation carries K to itself, and therefore it defines an automorphism of K.

***10.** Let K be a finite extension of a field F, and let $f(x) \in K[x]$. Prove that there is a nonzero polynomial $g(x) \in K[x]$ such that $f(x)g(x) \in F[x]$.

***11.** Let $f(x)$ be an irreducible quartic polynomial in $F[x]$. Let $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ be its roots in a splitting field K. Assume that the resolvent cubic has a root $\beta = \alpha_1\alpha_2 + \alpha_3\alpha_4$ in F, but that the discriminant D is not a square in F. According to the text, the Galois group of K/F is either C_4 or D_4.

(a) Determine the subgroup H of the group S_4 of permutations of the roots α_i which stabilizes β explicitly. Don't forget to prove that no permutations other than those you list fix β.

(b) Let $\gamma = \alpha_1\alpha_2 - \alpha_3\alpha_4$ and $\epsilon = \alpha_1 + \alpha_2 - \alpha_3 - \alpha_4$. Describe the action of H on these elements.

(c) Prove that γ^2 and ϵ^2 are in F.

(d) Let δ be the square root of the discriminant. Prove that if $\gamma \neq 0$, then $\delta\gamma$ is a square in F if and only if $G = C_4$. Similarly, prove that if $\epsilon \neq 0$, then $\delta\epsilon$ is a square in F if and only if $G = C_4$.

(e) Prove that γ and ϵ can't both be zero.

***12.** Let $F = \mathbb{F}_p(u, v)$ be a rational function field in two variables over the field $\mathbb{F}_p$ with p elements, and let $K = F(\alpha, \beta)$, where α, β are roots of the polynomials $x^p - u$ and $x^p - v$ respectively. Prove the following.

(a) The extension K/F has no primitive element.

(b) The elements $\gamma = \beta + c\alpha$, where $c \in F$, generate infinitely many different intermediate fields L.

***13.** Let K be a field with p^r elements. Prove that the Frobenius map defined by $\varphi(x) = x^p$ is a linear transformation of K, when K is viewed as a vector space the prime field $F = \mathbb{F}_p$, and determine its eigenvectors and eigenvalues.

Wie weit diese Methoden reichen werden, muss erst die Zukunft zeigen.

Emmy Noether

Appendix

Background Material

Historically speaking, it is of course quite untrue
that mathematics is free from contradiction;
non-contradiction appears as a goal to be achieved,
not as a God-given quality that has been granted us once for all.

Nicolas Bourbaki

1. SET THEORY

This section reviews some conventions about set theory which are used in this book, as well as some facts which will be referred to occasionally.

First, a remark about definitions: Any definition of a word or a phrase will have roughly the form

(1.1) *xxx* if @#&$% ,

where *xxx* is the word which is being defined and @#&$% is its defining property. For example, the sentence "An integer n is *positive* if $n > 0$" defines the notion of a positive integer. In a definition, the word *if* means *if and only if*. So in the definition of the positive integers, all integers which don't satisfy the requirement $n > 0$ are ruled out.

The notation

(1.2) {s ∈ S | @#&$%}

stands for the subset of S consisting of all elements s such that @#&$% is true. Thus if $\mathbb{Z}$ denotes the set of all integers, then $\mathbb{N} = \{n \in \mathbb{Z} \mid n > 0\}$ describes $\mathbb{N}$ as the set of positive integers or *natural numbers*.

Elements $a_1, \dots, a_n$ of a set are said to be *distinct* if no two of them are equal.

A *map* φ from a set S to a set T is any function whose *domain* of definition is S and whose *range* is T. The words *function* and *map* are used synonymously. We require that a function be single-valued. This means that every element $s \in S$ must have a uniquely determined *image* $\varphi(s) \in T$. The range T of φ is not required to be

the set of values of the function. By definition of a function, every image element $\varphi(s)$ is contained in T, but we allow the possibility that some elements $t \in T$ are not taken on by the function at all. We also take the domain and range of a function as part of its definition. If we restrict the domain to a subset, or if we extend the range, then the function obtained is considered to be different.

The domain and range of a map may also be described by the use of an arrow. Thus the notation $\varphi: S \rightarrow T$ tells us that φ is a map from S to T. The statement that $t = \varphi(s)$ may be described by a wiggly arrow: $s \rightsquigarrow t$ means that the element $s \in S$ is sent to $t \in T$ by the map under consideration. For example, the map $\varphi: \mathbb{Z} \rightarrow \mathbb{Z}$ such that $\varphi(n) = 2n + 1$ is described by $n \rightsquigarrow 2n + 1$.

The *image* of the map φ is the subset of T of elements which have the form $\varphi(s)$ for some $s \in S$. It will often be denoted by im φ, or by $\varphi(S)$:

$$\text{im } \varphi = \{t \in T \mid t = \varphi(s) \text{ for some } s \in S\}. \tag{1.3}$$

In case im φ is the whole range T, the map is said to be *surjective*. Thus φ is surjective if every $t \in T$ has the form $\varphi(s)$ for some $s \in S$.

The map φ is called *injective* if distinct elements s_1, s_2 of S have distinct images, that is, if $s_1 \neq s_2$ implies that $\varphi(s_1) \neq \varphi(s_2)$. A map which is both injective and surjective is called a *bijective* map. A *permutation* of a set S is a bijective map from S to itself.

Let $\varphi: S \rightarrow T$ and $\psi: T \rightarrow S$ be two maps. Then ψ is called an *inverse function* of φ if both of the composed maps $\varphi \circ \psi: T \rightarrow T$ and $\psi \circ \varphi: S \rightarrow S$ are the identity maps, that is, if $\varphi(\psi(t)) = t$ for all $t \in T$ and $\psi(\varphi(s)) = s$ for all $s \in S$. The inverse function is often denoted by φ^{-1}.

(1.4) **Proposition.** A map $\varphi: S \rightarrow T$ has an inverse function if and only if it is bijective.

Proof. Assume that φ has an inverse function ψ, and let us show that φ is both surjective and injective. Let t be any element of T, and let $s = \psi(t)$. Then $\varphi(s) = \varphi(\psi(t)) = t$. So t is in the image of φ. This shows that φ is surjective. Next, let s_1, s_2 be distinct elements of S, and let $t_i = \varphi(s_i)$. Then $\psi(t_i) = s_i$. So t_1, t_2 have distinct images in S, which shows that they are distinct. Therefore φ is injective. Conversely, assume that φ is bijective. Then since φ is surjective, every element $t \in T$ has the form $t = \varphi(s)$ for some $s \in S$. Since φ is injective, there can be only one such element s. So we define ψ by the following rule: $\psi(t)$ is the unique element $s \in S$ such that $\psi(s) = t$. This map is the required inverse function. □

Let $\varphi: S \rightarrow T$ be a map, and let U be a subset of T. The *inverse image* of U is defined to be the set

$$\varphi^{-1}(U) = \{s \in S \mid \varphi(s) \in U\}. \tag{1.5}$$

This set is defined whether or not φ has an inverse function. The notation φ^{-1}, as used here, is symbolic.

A set is called *finite* if it contains finitely many elements. If so, the number of its elements, sometimes called its *cardinality*, will be denoted by $|S|$. We will also

call this number the *order* of S. If S is infinite, we write $|S| = \infty$. The following theorem is quite elementary, but it is a very important principle.

(1.6) **Theorem.** Let $\varphi: S \to T$ be a map between finite sets.

(a) If φ is injective, then $|S| \leq |T|$.
(b) If φ is surjective, then $|S| \geq |T|$.
(c) If $|S| = |T|$, then φ is bijective if and only if it is either injective or surjective. □

The contrapositive of part (a) is often called the *pigeonhole principle*: *If* $|S| > |T|$, *then* φ *is not injective*. For example, if there are 87 socks in 79 drawers, then some drawer contains at least two socks.

An infinite set S is called *countable* if there is a bijective map $\varphi: \mathbb{N} \to S$ from the set of natural numbers to S. If there is no such map, then S is said to be *uncountable*.

(1.7) **Proposition.** The set $\mathbb{R}$ of real numbers is uncountable.

Proof. This proof is often referred to as Cantor's diagonal argument. Let $\varphi: \mathbb{N} \to \mathbb{R}$ be any map. We list the elements of the image of φ in the order $\varphi(1)$, $\varphi(2)$, $\varphi(3), \ldots$, and we write each of these real numbers in decimal notation. For example, the list might begin as follows:

$$
\begin{aligned}
\varphi(1) &= 8\ 2\ .\underline{3}\ 5\ 4\ 7\ 0\ 9\ 8\ 4\ 5\ 3\ 4 \ldots \\
\varphi(2) &= \quad\ \ .1\ \underline{2}\ 3\ 9\ 0\ 3\ 4\ 5\ 7\ 0\ 0 \ldots \\
\varphi(3) &= \quad 5\ .9\ 0\ \underline{8}\ 4\ 0\ 5\ 9\ 8\ 6\ 7\ 5 \ldots \\
\varphi(4) &= 1\ 2\ .8\ 7\ 4\ \underline{3}\ 5\ 2\ 6\ 4\ 4\ 4\ 4 \ldots \\
\varphi(5) &= \quad\ \ .0\ 0\ 1\ 4\ \underline{4}\ 1\ 0\ 0\ 3\ 4\ 9 \ldots \\
&\ \ \vdots \qquad\qquad \vdots
\end{aligned}
$$

We will now determine a real number which is not on the list. Consider the real number u whose decimal expansion consists of the underlined digits: $u = .3\ 2\ 8\ 3\ 4 \ldots$. We form a new real number by changing each of these digits, say

$$v = .4\ 5\ 1\ 4\ 2 \ldots .$$

Notice that $v \neq \varphi(1)$, because the first digit, 4, of v is not equal to the corresponding digit, 3, of $\varphi(1)$. Also, $v \neq \varphi(2)$, because the second digit, 5, of v is not equal to the corresponding digit of $\varphi(2)$. Similarly, $v \neq \varphi(n)$ for all n. This shows that φ is not surjective, which completes the proof, except for one point.

Some real numbers have two decimal expansions: .99999... is equal to 1.00000..., for example. This creates a problem with our argument. We have to choose v so that infinitely many of its digits are different from 9 and 0. The easiest way is to avoid these digits altogether. □

At a few places in the text, we refer to Zorn's Lemma, which is a tool for handling uncountable sets. We will now describe it. A *partial ordering* of a set S is a relation $s \leq s'$ which may hold between certain elements and which satisfies the following axioms for all s, s', s'' in S:

(1.8)

(i) $s \leq s$;
(ii) if $s \leq s'$ and $s' \leq s''$, then $s \leq s''$;
(iii) if $s \leq s'$ and $s' \leq s$, then $s = s'$.

A partial ordering is called a *total ordering* if in addition

(iv) for all s, s' in S, $s \leq s'$ or $s' \leq s$.

For example, let S be a set whose elements are sets. If A, B are in S, we may define $A \leq B$ if $A \subset B$. This is a partial ordering on S, called the *ordering by inclusion*. Whether or not it is a total ordering depends on the particular case.

If A is a subset of a partially ordered set S, then an *upper bound* for A is an element $b \in S$ such that for all $a \in A$, $a \leq b$. A partially ordered set S is called *inductive* if every totally ordered subset T of S has an upper bound in S.

A *maximal element* $m \in S$ is any element such that S contains no larger one, that is, such that there is no element $s \in S$ with $m \leq s$, except for m itself. This doesn't mean that m is an upper bound for S; in particular, there may be many different maximal elements. For example, the set of all proper subsets of $\{1,\dots,n\}$ contains n maximal elements, one of which is $\{1,3,4,\dots,n\}$.

(1.9) **Lemma.** *Zorn's Lemma:* An inductive partially ordered set has a maximal element. □

Zorn's Lemma is equivalent with the *axiom of choice*, which is known to be independent of the basic axioms of set theory. We will not enter into a further discussion of this equivalence, but we will show how Zorn's Lemma can be used to show that every vector space has a basis. Let us use unordered sets of vectors here.

(1.10) **Proposition.** Every vector space V over a field has a basis.

Proof. We take for S the set of (unordered) linearly independent subsets of V, partially ordered by inclusion, as above. We check that S is inductive: Let T be a totally ordered subset of S. Then we claim that the union of the sets making up T is also linearly independent; hence it is in S. To verify this, let

$$B = \bigcup_{A \in T} A$$

be the union. By definition, a relation of linear dependence on B is finite, so it can be written in the form

$$c_1 v_1 + \cdots + c_n v_n = 0, \tag{1.11}$$

with $v_i \in B$. Since B is a union of the sets $A \in T$, each v_i is contained in one of these subsets, call it A_i. Let i, j be two of the indices. Since T is totally ordered, $A_i \subset A_j$ or else $A_j \subset A_i$. It follows by induction that one of the sets, say A_i, contains all the others. Call this set A. Then $v_i \in A$ for all $i = 1,\ldots,n$. Since A is linearly independent, (1.11) is the trivial relation. This shows that B is linearly independent, hence that it is an element of S.

We have verified the hypothesis of Zorn's Lemma. So S contains a maximal element B, and we claim that B is a basis. By definition of S, B is linearly independent. Let $W = \text{Span}(B)$. If $W < V$, then we choose an element $v \in V$ which is not in W. Then the set $B \cup \{v\}$ is linearly independent [see Chapter 3 (3.10)]. This contradicts the maximality of B and shows that $W = V$, hence that B is a basis. □

A similar argument proves Theorem (8.3) of Chapter 10.

(1.12) **Proposition.** Let R be a ring. Every ideal $I \neq R$ is contained in a maximal ideal.

We leave this proof as an exercise. □

2. *TECHNIQUES OF PROOF*

Exactly what mathematicians consider an appropriate way to present a proof is not clearly defined. It isn't customary to give proofs which are complete in the sense that every step consists in applying a rule of logic to the previous step. Writing such a proof would take too long, and the main points wouldn't be emphasized. On the other hand, all difficult steps of the proof are supposed to be included. Someone reading the proof should be able to fill in as many details as needed to understand it. How to write a proof is a skill which can be learned only by experience.

We will discuss three important techniques used to construct proofs: *dichotomy, induction,* and *contradiction.*

The word *dichotomy* means division into two parts. It is used to subdivide a problem into smaller, more easily managed pieces. Other names for this procedure are *case analysis* and *divide and conquer*. Here is an example of dichotomy: One definition of the *binomial coefficient* $\binom{n}{k}$ (read *n choose k*) is that $\binom{n}{k}$ is the number of subsets of order k in the set $\{1, 2,\ldots, n\}$. For example, $\binom{4}{2} = 6$: The six subsets of order 2 of $\{1,2,3,4\}$ are $\{1,2\}$, $\{1,3\}$, $\{1,4\}$, $\{2,3\}$, $\{2,4\}$, $\{3,4\}$.

(2.1) **Proposition.** For every integer n and every $k \leq n$, $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$.

Proof. Let S be a subset of $\{1,2,\ldots,n\}$ of order k. Then either $n \in S$ or $n \notin S$. This is our dichotomy. If $n \notin S$, then S is actually a subset of $\{1, 2,\ldots, n-1\}$. By definition, there are $\binom{n-1}{k}$ of these subsets. Suppose that $n \in S$, and let $S' = S - \{n\}$ be the set obtained by deleting the element n from the set S. Then S' is a subset of $\{1, 2,\ldots, n-1\}$, of order $n - 1$. There are $\binom{n-1}{k-1}$ such sets S'. Hence there are $\binom{n-1}{k-1}$ subsets of order k which contain n. This gives us $\binom{n-1}{k} + \binom{n-1}{k-1}$ subsets of order k altogether. □

The remarkable power of the method of dichotomy is shown here: In each of the two cases, $n \in S$ and $n \notin S$, we have an additional fact about our set S. This additional fact can be used in the proof.

Often a proof will require sorting through several possibilities, examining each in turn. This is dichotomy, or case analysis. For instance, to determine the species of a plant, Gray's *Manual of Botany* leads through a sequence of dichotomies. A typical one is "leaves opposite on the stem (go to h), or leaves alternate (go to k)." Classification of mathematical structures will also proceed through a sequence of dichotomies. They need not be spelled out formally in simple cases, but when one is dealing with a complicated range of possibilities, careful sorting is needed. Here is a simple example:

(2.2) **Proposition.** Every group of order 4 is abelian.

Proof. Let G be a group of order 4, and let x, y be two elements of G. We are to show that $xy = yx$. Consider the five elements $1, x, y, xy, yx$. Since there are only four elements in the group, two of these must be equal. If $xy = yx$, the proposition is verified. We now run through the other possibilities:

Case 1: $x = 1$ or $y = 1$. If $x = 1$, then $xy = y = yx$. If $y = 1$, then $xy = x = yx$.

Case 2: $xy = 1$ or $yx = 1$. Then $y = x^{-1}$, and $xy = 1 = yx$.

Case 3: $x = y$. Then $xy = x^2 = yx$.

Case 4: Either $xy = x$, $yx = x$, $xy = y$, or $yx = y$. In the first two cases, we cancel x to conclude that $y = 1$, which puts us back in Case 1. In the last two cases, we cancel y.

This exhausts all possibilities and completes the proof. □

Induction is the main method for proving a sequence of statements P_n, indexed by positive integers n. To prove P_n for all n, the principle of induction requires us to do two things:

(2.3)

(i) prove that P_1 is true, and

(ii) prove that *if*, for some integer $k > 1$, P_k is true, *then* P_{k+1} is also true.

Sometimes it is more convenient to prove that if, for some integer $k \geq 0$, P_{k-1} is true, then P_k is true. This is just a change of the index.

Here are some examples of induction:

(2.4) **Proposition.** The determinant of an upper triangular matrix is the product of its diagonal entries.

Proof. Here P_n is the assertion that the proposition is true for an $n \times n$ triangular matrix. In case of a 1×1 matrix, there is only one diagonal entry, and it is

equal to the determinant. This means that P_1 is true. We now assume that P_{k-1} is true, and we prove P_k using that fact. Let A be a triangular $k \times k$ matrix. We expand the determinant by minors on the first column:

$$\det A = a_{11} \det A_{11} - a_{21} \det A_{21} + \cdots .$$

Since A is triangular the terms $a_{21}, a_{31}, \ldots, a_{k1}$ are all zero, so $\det A = a_{11} \det A_{11}$. Now notice that A_{11} is a $(k-1) \times (k-1)$ triangular matrix and that its diagonal entries are $a_{22}, a_{33}, \ldots, a_{kk}$. Since P_{k-1} is true by hypothesis, $\det A_{11}$ is the product $a_{22} \cdots a_{kk}$. Therefore $\det A = a_{11}a_{22} \cdots a_{kk}$, as required. □

(2.5) Proposition. $\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$.

Proof. Let P_r be the statement that $\binom{r}{k} = \dfrac{r!}{k!(n-k)!}$ for all $k = 1, \ldots, r$. Assume that P_{r-1} is true. Then the formula is true when we substitute $n = r - 1$ and $k = k$ and is also true when we substitute $n = r - 1$ and $k = k - 1$:

$$\binom{r-1}{k} = \frac{(r-1)!}{k!(r-1-k)!} \quad \text{and} \quad \binom{r-1}{k-1} = \frac{(r-1)!}{(k-1)!(r-k)!}.$$

According to Proposition (2.1), $\binom{r}{k} = \binom{r-1}{k} + \binom{r-1}{k-1}$. Thus

$$\binom{r}{k} = \binom{r-1}{k} + \binom{r-1}{k-1} = \frac{(r-1)!}{k!(r-1-k)!} + \frac{(r-1)!}{(k-1)!(r-k)!}$$

$$= \frac{r-k}{r} \frac{r!}{k!(r-k)!} + \frac{k}{r} \frac{r!}{k!(r-k)!} = \frac{r!}{k!(r-k)!}.$$

This shows that P_r is true, as required. □

As another example, let us prove the pigeonhole principle (1.6a), that if a map $\varphi: S \to T$ between finite sets is injective, then $|S| \le |T|$. We use induction on $n = |T|$. The assertion is true if $n = 0$, that is, if T is empty, because the only set which has a map to the empty set is the empty set.

We suppose that the theorem has been proved for $n = k - 1$, and we proceed to check it for $n = k$, where $k > 0$. We suppose that $|T| = k$, and we choose an element $t \in T$.

Case 1: t is in the image of φ. Since φ is injective, there is exactly one element $s \in S$ such that $\varphi(s) = t$. Let $S' = S - \{s\}$ and $T' = T - \{t\}$. Restricting φ to S', we obtain an injective map $\varphi': S' \to T'$. Since $|T'| = |T| - 1 = k - 1$, our induction hypothesis implies that $|S'| \le |T'|$. Therefore $|S| = |S'| + 1 \le |T'| + 1 = |T|$.

Case 2: t is not in im φ. In this case the image of φ is contained in $T' = T - \{t\}$. So φ defines an injective map $S \to T'$. Our induction hypothesis again implies that $|S| \le |T'| = |T| - 1$. □

There is a variant of the principle of induction, called *complete induction*. Here again, we wish to prove a statement P_n for each positive integer n. The principle of complete induction asserts that it is enough to prove the following statement:

(2.6) *If* n is a positive integer, and *if* P_k is true for every positive integer $k < n$, *then* P_n is true.

When $n = 1$, there are no positive integers $k < n$. So the hypothesis of (2.6) is automatically satisfied for $n = 1$. Hence a proof of (2.6) must include a proof of P_1.

The principle of complete induction is used when there is a procedure to reduce P_n to P_k for some smaller integers k, but not necessarily to P_{n-1}. Here is an example:

(2.7) **Theorem.** Every integer $n > 1$ is a product of prime integers.

An informal proof, which also exhibits an algorithm for finding a prime factorization, goes as follows: If n is a prime integer, then it is the product of one prime, and we are done. If not, then it has a divisor different from 1 and n. If n is given to us explicitly, we will be able to check whether or not there is such a proper divisor. If so, then n can be written as a product of integers, say $n = ab$, neither of which is 1, and then a and b are less than n. We continue factoring a and b if possible. Since the size of the factors decreases each time, this procedure can not be continued indefinitely, and eventually we end up with a prime factorization of n.

The principle of complete induction formalizes the statement that one can't continue replacing a positive integer by a smaller one infinitely often. To apply the principle, we let P_n be the statement that n is a product of primes, and we assume that P_k is true for all $k < n$. We go through the argument again. Either n is prime, in which case we are done, or else $n = ab$ and a and b are less than n. In this case the induction hypothesis tells us that P_a and P_b are both true, that is, that a and b are products of primes. Putting these products side by side gives us the required factorization of n.

The two proofs look slightly different from each other, because the algorithm is not mentioned in the statement of the theorem and has been partially suppressed in the formal proof. A better statement of the theorem would exhibit the algorithm:

(2.8) **Theorem.** The procedure of factoring an integer > 1 terminates after finitely many steps.

In this formulation, the formal proof becomes identical with the informal one. □

Proofs by contradiction proceed by assuming that the desired conclusion is false and deriving a contradiction from this assumption. The conclusion must therefore be true. We can, for example, rewrite the proof given above that a group of order 4 is abelian, in this way:

Proof of (2.2), Rewritten. We suppose that G is a nonabelian group of order 4, and we proceed to derive a contradiction from this assumption. Since G is not abelian, there are elements $x, y \in G$ such that $xy \neq yx$. Then y can not be any one of the elements $1, x, x^{-1}$, because those elements commute with x. Similarly, x is not equal to $1, y$, or y^{-1}. We may now check that the elements $1, x, y, xy, yx$ are distinct. This contradicts the hypothesis that $|G| = 4$. Therefore there does not exist a nonabelian group of order 4. □

Notice that there is no real difference between the two proofs of (2.2). The proof just given is really a fake contradiction argument, and, though logically correct, it is not aesthetically pleasing. One should avoid writing proofs in this way. On the other hand, there are true proofs by contradiction, in which the proof is not easily turned around to eliminate the contradiction. The proof given in the text [Chapter 6 (1.13)] that a group of order p^2, p a prime, is abelian is an example, as is the proof of (3.11) given below.

3. TOPOLOGY

This section reviews some concepts from topology which we will need from time to time. The sets which we want to study are subsets of Euclidean space $\mathbb{R}^k$.

Let r be a positive real number. The *open ball* of radius r about a point $X \in \mathbb{R}^k$ is the set of all points whose distance to X is less than r:

$$B_{X,r} = \{X' \in \mathbb{R}^k \mid |X' - X| < r\}. \tag{3.1}$$

A subset U of $\mathbb{R}^k$ is called *open* if whenever a point X lies in U the points sufficiently near to X also lie in U. In other words, U is open if it satisfies the following condition:

(3.2) *If $X \in U$ and if r is sufficiently small, then $B_{X,r} \subset U$.*

The radius r will depend on the point X.

Open sets have the following properties:

(3.3)

(i) The union of an arbitrary family of open sets is open.
(ii) The intersection of finitely many open sets is open.

The whole space $\mathbb{R}^k$ and the empty set $\emptyset$ are the simplest examples of open sets. Some more interesting open sets are obtained in this way: Let f be a continuous function $\mathbb{R}^k \to \mathbb{R}$. Then the sets

$$\{f > 0\}, \{f < 0\}, \{f \neq 0\} \tag{3.4}$$

are open. For instance, if $f(X) > 0$, then $f(X') > 0$ for all X' near X, because f is continuous. This shows that the general linear group $GL_2(\mathbb{R})$ is an open subset of the space $\mathbb{R}^4$ of all 2×2 matrices, because it is the set $\{\det P \neq 0\}$. Also, the open ball $B_{X,r}$ is an open set in $\mathbb{R}^k$, because it is defined by the inequality $|X' - X| - r < 0$.

Let S be any set in $\mathbb{R}^k$. We will also need the concept of open subset of S. By definition, a subset V of S is called *open* in S if whenever it contains a point X, then it also contains all points of S which are sufficiently near to X. This condition is explained by the following lemma:

(3.5) **Lemma.** Let V be a subset of a set S in $\mathbb{R}^k$. The following conditions on V are equivalent. If either one of them holds, then V is called an *open subset of* S:

(i) $V = U \cap S$ for some open set U of $\mathbb{R}^k$;
(ii) For every point $X \in V$, there is an $r > 0$ so that V contains the set $B_{X,r} \cap S$.

Proof. Assume that $V = U \cap S$ for some open set U of $\mathbb{R}^k$. Let $X \in V$. Then $X \in U$, and (3.2) guarantees the existence of an $r > 0$ such that $B_{X,r} \subset U$. So $B_{X,r} \cap S \subset U \cap S = V$, and (ii) is verified. Conversely, suppose that (ii) holds. For each $X \in V$, choose an open ball $B_{X,r}$ such that $B_{X,r} \cap S \subset V$, with the radius r depending as usual on the point X. Let U be the union of these balls. Then U is an open set in $\mathbb{R}^k$ (3.3i), and $U \cap S \subset V$. On the other hand, $X \in B_{X,r} \cap S \subset U \cap S$ for every $X \in V$. Therefore $V \subset U \cap S$, and $V = U \cap S$ as required. □

Open subsets of S have the properties (3.3), which follow from the same properties of open subsets of $\mathbb{R}^k$ because of (3.5i).

It is customary to speak of an open set V of S which contains a given point p as a *neighborhood* of p in S.

A subset C of a set S is called *closed* if its complement $(S - C)$ is open. For example, let f_i: $\mathbb{R}^k \to \mathbb{R}$ $(i = 1,\dots,k)$ be continuous functions. The locus

(3.6) $$\{f_1 = f_2 = \dots = f_k = 0\}$$

of solutions to the system of k equations $f_i = 0$ is a closed set in $\mathbb{R}^k$, because its complement is the union of the open sets $\{f_i \neq 0\}$. The 2-sphere $\{x_1^2 + x_2^2 + x_3^2 = 1\}$ is an example of a closed set in $\mathbb{R}^3$. So is the rotation group SO_2. It is the locus in $\mathbb{R}^{2\times 2}$ defined by the five equations

$$x_{11}x_{22} - x_{12}x_{21} = 1, \quad x_{11}{}^2 + x_{12}x_{21} = 1, \quad x_{21}x_{12} + x_{22}{}^2 = 1,$$

$$x_{11}x_{12} + x_{12}x_{22} = 0, \quad x_{21}x_{11} + x_{22}x_{21} = 0.$$

Closed sets have properties dual to (3.3):

(3.7)

(i) The intersection of an arbitrary family of closed sets is closed.
(ii) The union of finitely many closed sets is closed.

These rules follow from (3.3) by complementation.

A subset C of $\mathbb{R}^k$ is called *bounded* if the coordinates of the point in C are bounded, meaning that there is a positive real number b, a bound, such that for $X = (x_1,\dots,x_n) \in C$,

$$|x_i| \le b, \tag{3.8}$$

for all $i = 1,\dots,n$. If C is both closed and bounded, it is called a *compact* subset of $\mathbb{R}^k$. The unit 2-sphere is a compact set in $\mathbb{R}^3$.

Let S, T be subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$. A map $f: S \to T$ is called *continuous* if it carries nearby points of S to nearby points of T. Formally, the property of continuity is stated this way:

(3.9) *Let $s \in S$. For every real number $\epsilon > 0$, there is a $\delta > 0$ such that if $s' \in S$ and $|s' - s| < \delta$, then $|f(s') - f(s)| < \epsilon$.*

The easiest way to get a continuous map from S to T is as a restriction of a continuous map $F: \mathbb{R}^m \to \mathbb{R}^n$ which happens to carry S to T. Most of the maps we use are of this form. For example, the determinant is a continuous function from any one of the classical groups to $\mathbb{R}$ or $\mathbb{C}$.

A map $f: S \to S'$ is called a *homeomorphism* if it is bijective and if f^{-1}, as well as f, is continuous.

For example, the unit circle S^1 in $\mathbb{R}^2$ is homeomorphic to the rotation group SO_2. The homeomorphism $f: S^1 \to SO_2$ is given by restricting the map

$$F(x_1,x_2) = \begin{bmatrix} x_1 & x_2 \\ -x_2 & x_1 \end{bmatrix},$$

which carries $\mathbb{R}^2$ to the space $\mathbb{R}^4$ of 2×2 matrices. The map F is not bijective and is therefore not a homeomorphism, but it restricts to a homeomorphism f on the subsets S^1 and SO_2. Its inverse is the restriction to SO_2 of the projection $G: \mathbb{R}^4 \to \mathbb{R}^2$ which sends a 2×2 matrix to its top row. (The word *homeomorphism* must not be confused with *homomorphism*!)

A *path* is a continuous map $f: [0,1] \to \mathbb{R}^k$ from the unit interval to the space $\mathbb{R}^k$, and the path is said to lie in S if $f(t) \in S$ for every $t \in [0,1]$. A subset S of $\mathbb{R}^k$ is called *path-connected* if every pair of points p, $q \in S$ can be joined by a path lying in S. In other words, for every pair of points p, $q \in S$, there is a path f such that

(3.10)

(i) $f(t) \in S$ for all t in the interval;
(ii) $f(0) = p$ and $f(1) = q$.

Here is the most important property of path-connected sets:

(3.11) **Proposition.** A path-connected set S is not the disjoint union of proper open subsets. In other words, suppose that

$$S = \bigcup_{\alpha} V_i,$$

where V_i are open sets in S and $V_i \cap V_j = \emptyset$ if $i \neq j$. Then all but one of the sets V_i is empty.

Proof. Suppose that two of the sets are nonempty, say V_0 and V_1. We set aside V_0 and replace V_1 by the union of the remaining subsets, which is open by (3.3). Then $V_0 \cup V_1 = S$ and $V_0 \cap V_1 = \emptyset$. This reduces to the case that there are exactly two open sets.

Choose points $p \in V_0$ and $q \in V_1$, and let f: $[0, 1] \rightarrow S$ be a path in S connecting p to q. We will obtain a contradiction by examining the path at the point where it leaves V_0 for the last time.

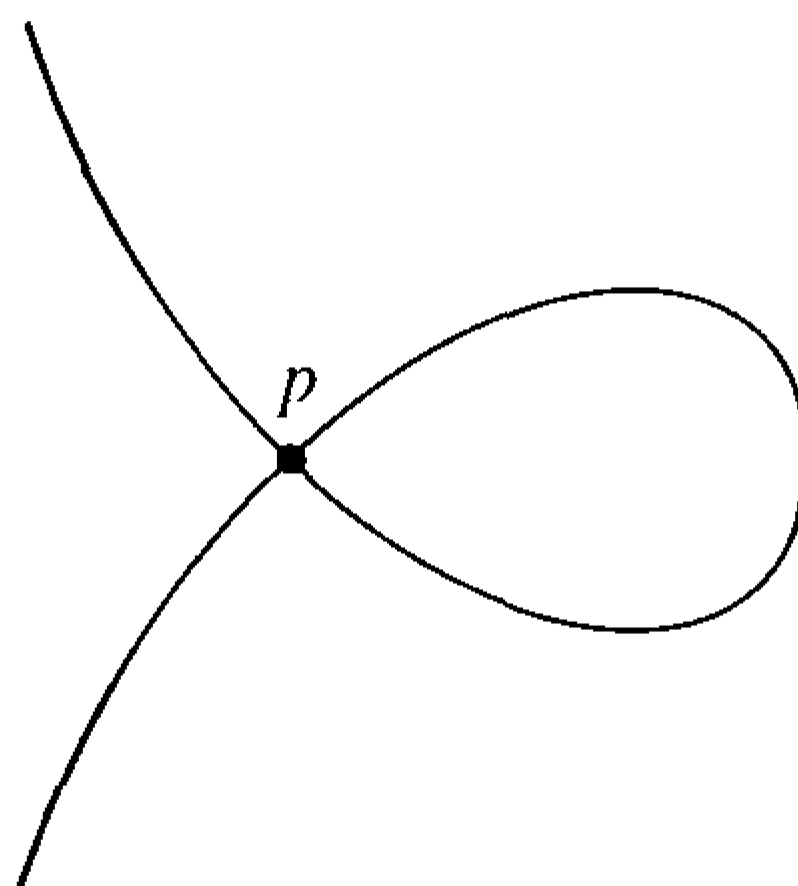
Let b be the least upper bound of all $t \in [0, 1]$ such that $f(t) \in V_0$, and let $X = f(b)$. If $X \in V_0$, then all points of $B_{X,r} \cap S$ are in V_0, if r is small enough. Since f is continuous, $f(t) \in B_{X,r}$ for all t sufficiently near b. So $f(t) \in V_0$ for these points. Taking t slightly larger than b contradicts the choice of b as an upper bound of the points mapping to V_0. Therefore X is not in V_0, so it has to be in V_1. But reasoning in the same way, we find that $f(t) \in V_1$ for all t sufficiently near b. Taking t slightly smaller than b contradicts the choice of b as the least upper bound of points mapping to V_0. This contradiction completes the proof. □

The final concept from topology is that of manifold.

(3.12) **Definition.** A subset S of $\mathbb{R}^n$ is called a *manifold of dimension d* if every point p of S has a neighborhood in S which is homeomorphic to an open set in $\mathbb{R}^d$.

For example, the sphere $\{(x, y, z) \mid x^2 + y^2 + z^2 = 1\}$ is a two-dimensional manifold. The half sphere $U = \{z > 0\}$ is open in S^3 (3.4, 3.5) and projects continuously to the unit ball $B_{0,1} = \{x_1^2 + x_2^2 + x_3^2 < 1\}$ in $\mathbb{R}^3$. The inverse function $z = \sqrt{1 - x^2 - y^2}$ is continuous. Therefore U is homeomorphic to $B_{0,1}$. Since the 3-sphere is covered by such half spheres, it is a manifold.

The figure below shows a set which is not a manifold. It becomes a manifold of dimension 1 when the point p is deleted. Note that homogeneity is false for this set. It looks different near p from how it looks near the other points.



(3.13) **Figure.** A set which is not a manifold.

4. THE IMPLICIT FUNCTION THEOREM

The Implicit Function Theorem is used at two places in this book, so we state it here for reference.

(4.1) **Theorem.** *Implicit Function Theorem:* Let $f(x,y) = (f_1(x,y),\dots,f_r(x,y))$ be functions of $n + r$ real variables $(x,y) = (x_1,\dots,x_m,y_1,\dots,y_r)$, which have continuous partial derivatives in an open set of $\mathbb{R}^{n+r}$ containing the point (a,b). Assume that the Jacobian determinant

$$\det \begin{bmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_r} \\ \vdots & & \vdots \\ \frac{\partial f_r}{\partial y_1} & \cdots & \frac{\partial f_r}{\partial y_r} \end{bmatrix}$$

is not zero at the point (a,b). There is a neighborhood U of the point a in $\mathbb{R}^n$ such that there are unique continuously differentiable functions $Y_1(x),\dots,Y_r(x)$ on U satisfying

$$f(x,Y(x)) = 0 \quad \text{and} \quad Y(a) = b.$$

The Implicit Function Theorem is closely related to the Inverse Function Theorem, which is used in Chapter 8 (5.8):

(4.2) **Theorem.** *Inverse Function Theorem:* Let f be a continuously differentiable map from an open set U of $\mathbb{R}^n$ to $\mathbb{R}^n$. Assume that the Jacobian determinant

$$\det \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

is not zero at a point $a \in \mathbb{R}^n$. There is a neighborhood of a on which f has a continuously differentiable inverse function.

We refer to the book by Rudin listed in the Suggestions for Further Reading for proofs of these two theorems. □

We also use the following complex analogue of the Implicit Function Theorem in one place [Chapter 13 (8.14)]:

(4.3) **Theorem.** Let $f(x,y)$ be a complex polynomial. Suppose that for some

$(a, b) \in \mathbb{C}^2$, we have $f(a, b) = 0$ and $\frac{\partial f}{\partial y}(a, b) \neq 0$. There is a neighborhood U of x in $\mathbb{C}$ on which a unique continuous function $Y(x)$ exists having the properties

$$f(x, Y(x)) = 0 \quad \text{and} \quad Y(a) = b.$$

Since references for this extension are not so common, we will give a proof which reduces it to the real Implicit Function Theorem. The method is simply to write everything in terms of its real and imaginary parts and then to verify the hypotheses of (4.1). The same argument will apply with more variables.

Proof. We write $x = x_0 + x_1 i$, $y = y_0 + y_1 i$, $f = f_0 + f_1 i$, where $f_i = f_i(x_0, x_1, y_0, y_1)$ is a real-valued function of four real variables. We are to solve the pair of equations $f_0 = f_1 = 0$ for y_0, y_1 as functions of x_0, x_1. According to (4.1), we have to prove that the Jacobian determinant

$$\det \begin{bmatrix} \frac{\partial f_0}{\partial y_0} & \frac{\partial f_0}{\partial y_1} \\ \frac{\partial f_1}{\partial y_0} & \frac{\partial f_1}{\partial y_1} \end{bmatrix}$$

is not zero at (a, b). Since f is a polynomial in x, y, the real functions f_i are also polynomials in x_i, y_j. So they have continuous derivatives.

(4.4) **Lemma.** Let $f(x,y)$ be a polynomial with complex coefficients. With the above notation,

(i) $\frac{\partial f}{\partial y} = \frac{\partial f_0}{\partial y_0} + \frac{\partial f_1}{\partial y_0} i$, and

(ii) the *Cauchy–Riemann equations* $\frac{\partial f_0}{\partial y_0} = \frac{\partial f_1}{\partial y_1}$ and $\frac{\partial f_0}{\partial y_1} = -\frac{\partial f_1}{\partial y_0}$ hold.

Proof of the Lemma. Since f is a polynomial and since the derivative of a sum is the sum of the derivatives, it is enough to prove the lemma for the monomials $cy^n = (c_0 + c_1 i)(y_0 + y_1 i)^n$. For these monomials, the lemma follows from the product rule for differentiation, by induction on n. □

We return to the proof of Theorem (4.3). By hypothesis, $f_i(a_0, a_1, b_0, b_1) = 0$. Also, since $\frac{\partial f}{\partial y}(a, b) \neq 0$, we know by (4.4i) that $\frac{\partial f_0}{\partial y_0} = d_0$ and $\frac{\partial f_1}{\partial y_0} = d_1$ are not both zero. By (4.4ii), the Jacobian determinant is

$$\det \begin{bmatrix} d_0 & -d_1 \\ d_1 & d_0 \end{bmatrix} = d_0^2 + d_1^2 > 0.$$

This shows that the hypotheses of the Implicit Function Theorem (4.1) are satisfied. □

EXERCISES

1. Set Theory

1. Let $\varphi: \mathbb{Z} \to \mathbb{R}$ be the map defined by $\varphi(n) = n^3 - 3n + 1$.
   (a) Is φ injective?
   (b) Determine $\varphi^{-1}(U)$, where U is the interval (i) $[0, \infty)$, (ii) $[2, 4]$, (iii) $[4, 12]$.
2. Give an example of a map $\varphi: S \to S$ from an infinite set to itself which is surjective but not injective, and one which is injective but not surjective.
3. Let $\varphi: S \to T$ be a map of sets.
   (a) Let U be a subset of S. Prove that $\varphi(\varphi^{-1}(U)) \subset U$ and that if φ is surjective, then $\varphi(\varphi^{-1}(U)) = U$.
   (b) Let V be a subset of T. Prove that $\varphi^{-1}(\varphi(V)) \supset V$ and that if φ is injective, then $\varphi^{-1}(\varphi(V)) = V$.
4. Let $\varphi: S \to T$ be a map of nonempty sets. A map $\psi: T \to S$ is a *left inverse* if $\psi \circ \varphi: S \to S$ is the identity and a *right inverse* if $\varphi \circ \psi: T \to T$ is the identity. Prove that φ has a left inverse if and only if it is injective and has a right inverse if and only if it is surjective.
5. Let S be a partially ordered set.
   (a) Prove that if S contains an upper bound b for S, then b is unique, and also b is a maximal element.
   (b) Prove that if S is totally ordered, then a maximal element m is an upper bound for S.
6. (a) Describe precisely which real numbers have more than one decimal expansion and how many expansions such a number has.
   (b) Fix the proof of Proposition (1.7).
7. Use Zorn's Lemma to prove that every ideal $I \neq R$ is contained in a maximal ideal. Do this by showing that the set S of all ideals $I \neq R$, ordered by inclusion, is inductive.

2. Techniques of Proof

1. Use induction to find a closed form for each of the following expressions.
   (a) $1 + 3 + 5 + \cdots + (2n + 1)$
   (b) $1^2 + 2^2 + 3^3 + \cdots + n^2$
   (c) $1 + 1/2 + 1/3 + \cdots + 1/n$
   (d) $\dfrac{1}{1 \cdot 2} + \dfrac{1}{2 \cdot 3} + \cdots + \dfrac{1}{n(n+1)}$
2. Prove that $1^3 + 2^3 + \cdots + n^3 = (n(n + 1))^2/4$.
3. Prove that $1/(1 \cdot 2) + 1/(2 \cdot 3) + \cdots + 1/(n(n + 1)) = n/(n + 1)$.
4. Let S, T be finite sets.
   (a) Let $\varphi: S \longrightarrow T$ be an injective map. Prove by induction that $|S| \leq |T|$ and that if $|S| = |T|$, then φ is bijective.
   (b) Let $\varphi: S \longrightarrow T$ be a surjective map. Prove by induction that $|S| \geq |T|$ and that if $|S| = |T|$, then φ is bijective.
5. Let n be a positive integer. Show that if $2^n - 1$ is a prime number, then n is prime.

6. Let $a_n = 2^{2^n} + 1$. Prove that $a_n = a_0 a_1 \cdots a_{n-1} + 2$.
7. A polynomial with rational coefficients is called irreducible if it is not constant and if it is not a product of two nonconstant polynomials whose coefficients are rational numbers. Use complete induction to prove that every polynomial with rational coefficients can be written as a product of irreducible polynomials.
8. Prove parts (b) and (c) of Theorem (1.6).

3. Topology

1. Let S be a subset of $\mathbb{R}^k$, and let f, g be continuous functions from S to $\mathbb{R}$. Determine whether or not the following subsets are open or closed in S.
   (a) $\{f(X) \geq 0\}$ **(b)** $\{f(X) \neq 2\}$ **(c)** $\{f(X) < 0, g(X) > 0\}$
   (d) $\{f(X) \leq 0, g(X) < 0\}$ **(e)** $\{f(X) \neq 0, g(X) = 0\}$ **(f)** $\{f(X) \in \mathbb{Z}\}$
   (g) $\{f(X) \in \mathbb{Q}\}$
2. Let $X \in \mathbb{R}^n$. Determine whether or not the following sets are open or closed.
   (a) $\{rX \mid r \in \mathbb{R}, r > 0\}$ **(b)** $\{rX \mid r \in \mathbb{R}, r \geq 0\}$
3. **(a)** Let $P = (p_{ij})$ be an invertible matrix, and let $d = \det P$. We can define a map $GL_n(\mathbb{R}) \longrightarrow \mathbb{R}^{n^2+1}$ by sending $P \rightsquigarrow (p_{ij}, d)$. Show that this rule embeds $GL_n(\mathbb{R})$ as a closed set in $\mathbb{R}^{n^2+1}$.
   (b) Illustrate this map in the case of $GL_1(\mathbb{R})$.
4. Prove that the product of $M \times M'$ two manifolds M, M' is a manifold.
5. Show that $SL_2(\mathbb{R})$ is not a compact group.
6. **(a)** Sketch the curve C: $x_2^2 = x_1^3 - x_1^2$ in $\mathbb{R}^2$.
   (b) Prove that this locus is a manifold of dimension 1 if the origin is deleted.

4. The Implicit Function Theorem

1. Prove Lemma (4.4).
2. Prove that $SL_2(\mathbb{R})$ is a manifold, and determine its dimension.
3. Let $f(x, y)$ be a complex polynomial. Assume that the equations

$$f = 0, \quad \frac{\partial f}{\partial x} = 0, \quad \frac{\partial f}{\partial y} = 0$$

have no common solution in $\mathbb{C}^2$. Prove that the locus $f = 0$ is a manifold of dimension 2.

NOTATION

A_n	the *alternating group*, Chapter 2 (4.7)
$B_{x,r}$	the *open ball* of radius r about the point X, Appendix (3.1)
$\mathbb{C}$	the field of *complex numbers*, Chapter 2 (1.11)
C_n	the *cyclic group* of order n, Chapter 5 (3.4)
D_n	the *dihedral group*, Chapter 5 (3.4)
det	*determinant*, Chapter 1 (3.4)
$\mathbb{F}_p$	the *prime field* $\mathbb{Z}/(p)$, Chapter 3 (2.4)
GL_n	the *general linear group*, Chapter 2 (1.13)
I	the *identity matrix*, Chapter 1 (1.14)
I	the *icosahedral group*, Chapter 5 (9.1)
im φ	the *image* of the map φ, Appendix (1.3)
ker φ	the *kernel* of the homomorphism φ, Chapter 2 (4.5)
ℓ^∞	the space of *bounded sequences*, Chapter 3 (5.2)
M	the *group of motions* of the plane, Chapter 4 (5.15), Chapter 5 (2.1)
$N(H)$	the *normalizer* of H, Chapter 6 (3.7)
$\mathbb{N}$	the set of positive integers, or *natural numbers*, Chapter 10 (2.1)
O_n	the *orthogonal group*, Chapter 5 (5.3), Chapter 8 (1.3)
$O_{3,1}$	the *Lorentz group*, Chapter 8 (1.4)
PSL_n	the *projective group*, Chapter 8 (8.2)
$\mathbb{R}$	the field of *real numbers*, Chapter 2 (1.11)
$\mathbb{R}^n$	the space of *n-dimensional vectors*, Chapter 3 (1.1)
S_n	the *symmetric group*, Chapter 2 (1.14)
S^n	the *n-dimensional sphere*, Chapter 8 (2.6)
SL_n	the *special linear group*, Chapter 2 (4.6), Chapter 8 (1.8)
SO_n	the *special orthogonal group*, Chapter 4 (5.4), Chapter 8 (1.8)
SP_{2n}	the *symplectic group*, Chapter 8 (1.6)
SU_n	the *special unitary group*, Chapter 8 (1.8)
T	the *tetrahedral group*, Chapter 5 (9.1)
^{t}	(superscript t) the *transpose* of a matrix, Chapter 1 (2.24)
tr	*trace*, Chapter 4 (4.18)
U_n	the *unitary group*, Chapter 7 (4.15), Chapter 8 (1.8)
Z	the *center of a group*, Chapter 2 (4.10)
$\mathbb{Z}$	the ring of *integers*, Chapter 2 (1.11)
$Z(x)$	the *centralizer* of x, Chapter 6 (1.5)
$*$	If A is a complex matrix, then $A^* = \overline{A}^t$, Chapter 7 (4.7) In a matrix display, $*$ denotes an undetermined entry, Chapter 1 (1.15) The starred exercises are some of the more difficult ones.
$^+$	(superscript $+$) The group whose law of composition is *addition*, Chapter 2 (1.1)
$^\times$	(superscript $\times$) The group whose law of composition is *multiplication*, Chapter 2 (1.1)
$\oplus$	*direct sum*, Chapter 3 (6.4), Chapter 12 (6.3)
!	*factorial; n*! is the product of the integers $1, 2, \ldots, n$.

$\binom{n}{k}$ a *binomial coefficient*, Appendix (2.1)
$[\mu]$ the largest integer $\leq \mu$, Chapter 11 (10.23)

If S and T are sets, we use the following notation:

$|S|$ the number of elements, also called the *order* of the set S.
$s \in S$ s is an *element of* S.
$S \subset T$ S is a *subset* of T, or S is *contained in* T. In other words, every element of S is also an element of T.
$T \supset S$ T *contains* S, which is the same as $S \subset T$.
$S < T$ S is a *proper subset* of T, meaning that it is a subset, and that T contains an element which is not a member of S.
$T > S$ This is the same as $S < T$.
$T - S$ This notation is used only when S is a subset of T, and then it denotes the *complement* of S in T, the set of all elements which are in T but not in S:

$$T - S = \{x \mid x \in T \text{ but } x \notin S\}.$$

$S \cap T$ The *intersection* of the sets S and T, which is the set of all elements in common to S and T.
$S \cup T$ The *union* of the sets S and T, which is the set of all elements x which are contained in at least one of the sets S and T.
$S \times T$ the *product set*. Its elements are ordered pairs (s, t) of elements:

$$S \times T = \{(s, t) \mid s \in S, t \in T\}.$$

Since the parentheses have other meanings, we sometimes leave them off, and denote an element of the product set by s,t.
$\varphi: S \to T$ a *map* φ from S to T, or a *function* whose domain is S and whose range is T.
$s \rightsquigarrow t$ The wiggly arrow indicates that the map under consideration sends the element s to the element t, i.e., that $\varphi(s) = t$.
□ This symbol indicates that a digression in the text, such as a proof or an example, has ended, and that the text returns to the main thread. □

Suggestions for Further Reading

General algebra texts:

G. Birkhoff and S. MacLane, *A Survey of Modern Algebra*, 3rd ed, Macmillan, New York, 1965.

I. N. Herstein, *Topics in Algebra,* 2nd ed, Wiley, New York, 1975.

N. Jacobson, *Basic Algebra I,II,* Freeman, San Francisco, 1974, 1980.

S. Lang, *Algebra*, 2nd ed, Addison-Wesley, Reading, MA, 1965.

H. Paley and P. M. Weichsel, *Elements of Abstract and Linear Algebra*, Holt, Reinhart and Winston, New York, 1972.

B. L. van der Waerden, *Modern Algebra*, Ungar, New York, 1970.

History of mathematics:

N. Bourbaki, *Eléments d'histoire des mathématiques*, Hermann, Paris, 1974.

H. Edwards, *Fermat's Last Theorem*, Springer-Verlag, New York, 1977.

H. Edwards, *Galois Theory*, Springer-Verlag, New York, 1984.

Morris Klein, *Mathematical Thought from Ancient to Modern Times,* Oxford, New York, 1972.

B. L. van der Waerden, *A History of Algebra*, Springer-Verlag, Berlin, New York, 1985.

Chapter 1:

T. Muir, *A Treatise on the Theory of Determinants*, Dover, New York, 1960.

Chapter 2:

I. N. Herstein, *Topics in Algebra*, 2nd ed, Wiley, New York, 1975.

Chapters 3,4:

G. Strang, *Linear Algebra and Its Applications*, 3rd ed, Harcourt Brace Jovanovich, San Diego, 1988.

Chapter 5:

C. T. Benson and L. C. Grove, *Finite Reflection Groups*, 2nd ed., Springer-Verlag, New York, 1985.

H. M. S. Coxeter, *Introduction to Geometry*, Wiley, New York, 1961.

L. Ford, *Automorphic Functions*, Chelsea, New York, 1929.

B. Grünbaum and G. C. Sheppard, *Tilings and Patterns*, W. H. Freeman, New York, 1967.

H. W. Guggenheimer, *Plane Geometry and Its Groups*, Holden-Day, San Francisco, 1967.

Chapter 7:

B. Noble, *Applied Linear Algebra*, 2nd ed., Prentice Hall, Englewood Cliffs, NJ, 1977.

Chapter 8:

R. Howe, "Very Basic Lie Theory," *Math Monthly* 90 (1983) 600–623.

F. Warner, *Foundations of Differential Geometry and Lie Groups*, Springer-Verlag, New York, 1983.

H. Weyl, *The Classical Groups*, Princeton University Press, Princeton, 1946.

Chapter 9:

J.-P. Serre, *Linear Representations of Finite Groups*, Springer-Verlag, New York, 1977.

Chapter 10:

K. Kendig, *Elementary Algebraic Geometry*, Springer-Verlag, New York, 1976.

E. Landau, *Foundations of Analysis*, Chelsea, New York, 1960.

Chapter 11:

Z. I. Borevich and I. R. Shafarevitch, *Number Theory*, Academic Press, New York, 1966.

H. Edwards, *Fermat's Last Theorem*, Springer-Verlag, New York, 1977.

K. F. Gauss, *Disquisitiones Arithmeticae*, Leipzig, 1801.

H. Hasse, *Number Theory*, Springer-Verlag, New York, 1980.

J.-P. Serre, *A Course in Arithmetic*, Springer-Verlag, New York, 1973.

H. Stark, *An Introduction to Number Theory*, M.I.T. Press, Cambridge, 1978.

Chapter 13:

G. A. Bliss, *Algebraic Functions*, AMS Colloquium Publications No. XVI, New York, 1933.

Chapter 14:

H. Edwards, *Galois Theory*, Springer-Verlag, New York, 1984.

Appendix:

J. R. Munkres, *Topology; A First Course*, Prentice Hall, Englewood Cliffs, NJ, 1975.

W. Rudin, *Principles of Mathematical Analysis*, 3rd ed, McGraw-Hill, New York, 1976.

Index